1, TITLE: Dual Cross-Attention Learning for Fine-Grained Visual Categorization and Object Re-Identification
http://openaccess.thecvf.com//content/CVPR2022/html/Zhu_Dual_Cross-Attention_Learning_for_Fine-Grained_Visual_Categorization_and_Object_Re-Identification_CVPR_2022_paper.html
AUTHORS: Haowei Zhu, Wenjing Ke, Dong Li, Ji Liu, Lu Tian, Yi Shan
HIGHLIGHT: In this work, we explore how to extend self-attention modules to better learn subtle feature embeddings for recognizing fine-grained objects, e.g., different bird species or person identities.

2, TITLE: SimAN: Exploring Self-Supervised Representation Learning of Scene Text via Similarity-Aware Normalization
http://openaccess.thecvf.com//content/CVPR2022/html/Luo_SimAN_Exploring_Self-Supervised_Representation_Learning_of_Scene_Text_via_Similarity-Aware_CVPR_2022_paper.html
AUTHORS: Canjie Luo, Lianwen Jin, Jingdong Chen
HIGHLIGHT: Specifically, we propose a Similarity-Aware Normalization (SimAN) module to identify the different patterns and align the corresponding styles from the guiding patch.

3, TITLE: GASP, a Generalized Framework for Agglomerative Clustering of Signed Graphs and Its Application to Instance Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Bailoni_GASP_a_Generalized_Framework_for_Agglomerative_Clustering_of_Signed_Graphs_CVPR_2022_paper.html
AUTHORS: Alberto Bailoni, Constantin Pape, Nathan Hütsch, Steffen Wolf, Thorsten Beier, Anna Kreshuk, Fred A. Hamprecht
HIGHLIGHT: We propose a theoretical framework that generalizes simple and fast algorithms for hierarchical agglomerative clustering to weighted graphs with both attractive and repulsive interactions between the nodes.

4, TITLE: Estimating Example Difficulty Using Variance of Gradients
http://openaccess.thecvf.com//content/CVPR2022/html/Agarwal_Estimating_Example_Difficulty_Using_Variance_of_Gradients_CVPR_2022_paper.html
AUTHORS: Chirag Agarwal, Daniel D'souza, Sara Hooker
HIGHLIGHT: In this work, we propose Variance of Gradients (VoG) as a valuable and efficient metric to rank data by difficulty and to surface a tractable subset of the most challenging examples for human-in-the-loop auditing.

5, TITLE: One Loss for Quantization: Deep Hashing With Discrete Wasserstein Distributional Matching
http://openaccess.thecvf.com//content/CVPR2022/html/Doan_One_Loss_for_Quantization_Deep_Hashing_With_Discrete_Wasserstein_Distributional_CVPR_2022_paper.html
AUTHORS: Khoa D. Doan, Peng Yang, Ping Li
HIGHLIGHT: This paper considers an alternative approach to learning the quantization constraints.

6, TITLE: Pixel Screening Based Intermediate Correction for Blind Deblurring
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Pixel_Screening_Based_Intermediate_Correction_for_Blind_Deblurring_CVPR_2022_paper.html
AUTHORS: Meina Zhang, Yingying Fang, Guoxi Ni, Tieyong Zeng
HIGHLIGHT: However, these methods still fail while dealing with images containing saturations and large blurs. To address this problem, we propose an intermediate image correction method which utilizes Bayes posterior estimation to screen through the intermediate image and exclude those unfavorable pixels to reduce their influence for kernel estimation.

7, TITLE: Weakly Supervised Semantic Segmentation by Pixel-to-Prototype Contrast
http://openaccess.thecvf.com//content/CVPR2022/html/Du_Weakly_Supervised_Semantic_Segmentation_by_Pixel-to-Prototype_Contrast_CVPR_2022_paper.html
AUTHORS: Ye Du, Zehua Fu, Qingjie Liu, Yunhong Wang
HIGHLIGHT: In this study, we propose weakly-supervised pixel-to-prototype contrast that can provide pixel-level supervisory signals to narrow the gap.

8, TITLE: Controllable Animation of Fluid Elements in Still Images
http://openaccess.thecvf.com//content/CVPR2022/html/Mahapatra_Controllable_Animation_of_Fluid_Elements_in_Still_Images_CVPR_2022_paper.html
AUTHORS: Aniruddha Mahapatra, Kuldeep Kulkarni
HIGHLIGHT: We propose a method to interactively control the animation of fluid elements in still images to generate cinemagraphs.

9, TITLE: Holocurtains: Programming Light Curtains via Binary Holography
http://openaccess.thecvf.com//content/CVPR2022/html/Chan_Holocurtains_Programming_Light_Curtains_via_Binary_Holography_CVPR_2022_paper.html
AUTHORS: Dorian Chan, Srinivasa G. Narasimhan, Matthew O'Toole
HIGHLIGHT: In this work, we propose Holocurtains: a light-efficient approach to producing light curtains of arbitrary shape.

10, TITLE: Recurrent Dynamic Embedding for Video Object Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Recurrent_Dynamic_Embedding_for_Video_Object_Segmentation_CVPR_2022_paper.html
AUTHORS: Mingxing Li, Li Hu, Zhiwei Xiong, Bang Zhang, Pan Pan, Dong Liu
HIGHLIGHT: In this paper, we propose a Recurrent Dynamic Embedding (RDE) to build a memory bank of constant size.

11, TITLE:		Deep Hierarchical Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Deep_Hierarchical_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS:		Liulei Li, Tianfei Zhou, Wenguan Wang, Jianwu Li, Yi Yang
HIGHLIGHT:		In this paper, we instead address hierarchical semantic segmentation (HSS), which aims at structured, pixel-wise description of visual observation in terms of a class hierarchy.

12, TITLE:		f-SfT: Shape-From-Template With a Physics-Based Deformation Model
http://openaccess.thecvf.com//content/CVPR2022/html/Kairanda_f-SfT_Shape-From-Template_With_a_Physics-Based_Deformation_Model_CVPR_2022_paper.html
AUTHORS:		Navami Kairanda, Edith Tretschk, Mohamed Elgharib, Christian Theobalt, Vladislav Golyanik
HIGHLIGHT:		In contrast to previous works, this paper proposes a new SfT approach explaining 2D observations through physical simulations accounting for forces and material properties.

13, TITLE:		Continual Object Detection via Prototypical Task Correlation Guided Gating Mechanism
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Continual_Object_Detection_via_Prototypical_Task_Correlation_Guided_Gating_Mechanism_CVPR_2022_paper.html
AUTHORS:		Binbin Yang, Xinchi Deng, Han Shi, Changlin Li, Gengwei Zhang, Hang Xu, Shen Zhao, Liang Lin, Xiaodan Liang
HIGHLIGHT:		Different from previous works that tune the whole network for all tasks, in this work, we present a simple and flexible framework for continual object detection via pRotOtypical taSk corrElaTion guided gaTing mechAnism (ROSETTA).

14, TITLE:		DATA: Domain-Aware and Task-Aware Self-Supervised Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Chang_DATA_Domain-Aware_and_Task-Aware_Self-Supervised_Learning_CVPR_2022_paper.html
AUTHORS:		Qing Chang, Junran Peng, Lingxi Xie, Jiajun Sun, Haoran Yin, Qi Tian, Zhaoxiang Zhang
HIGHLIGHT:		In this paper, we present DATA, a simple yet effective NAS approach specialized for SSL that provides Domain-Aware and Task-Aware pre-training.

15, TITLE:		TWIST: Two-Way Inter-Label Self-Training for Semi-Supervised 3D Instance Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Chu_TWIST_Two-Way_Inter-Label_Self-Training_for_Semi-Supervised_3D_Instance_Segmentation_CVPR_2022_paper.html
AUTHORS:		Ruihang Chu, Xiaoqing Ye, Zhengzhe Liu, Xiao Tan, Xiaojuan Qi, Chi-Wing Fu, Jiaya Jia
HIGHLIGHT:		To leverage the unlabeled data to boost model performance, we present a novel Two-Way Inter-label Self-Training framework named TWIST.

16, TITLE:		Voxel Set Transformer: A Set-to-Set Approach to 3D Object Detection From Point Clouds
http://openaccess.thecvf.com//content/CVPR2022/html/He_Voxel_Set_Transformer_A_Set-to-Set_Approach_to_3D_Object_Detection_CVPR_2022_paper.html
AUTHORS:		Chenhang He, Ruihuang Li, Shuai Li, Lei Zhang
HIGHLIGHT:		In this paper, we propose a novel voxel-based architecture, namely Voxel Set Transformer (VoxSeT), to detect 3D objects from point clouds by means of set-to-set translation.

17, TITLE:		Learning Adaptive Warping for Real-World Rolling Shutter Correction
http://openaccess.thecvf.com//content/CVPR2022/html/Cao_Learning_Adaptive_Warping_for_Real-World_Rolling_Shutter_Correction_CVPR_2022_paper.html
AUTHORS:		Mingdeng Cao, Zhihang Zhong, Jiahao Wang, Yinqiang Zheng, Yujiu Yang
HIGHLIGHT:		This paper proposes a real-world rolling shutter (RS) correction dataset, BS-RSC, and a corresponding model to correct the RS frames in a distorted video.

18, TITLE:		Siamese Contrastive Embedding Network for Compositional Zero-Shot Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Siamese_Contrastive_Embedding_Network_for_Compositional_Zero-Shot_Learning_CVPR_2022_paper.html
AUTHORS:		Xiangyu Li, Xu Yang, Kun Wei, Cheng Deng, Muli Yang
HIGHLIGHT:		In this paper, we propose a novel Siamese Contrastive Embedding Network (SCEN) for unseen composition recognition.

19, TITLE:		Bongard-HOI: Benchmarking Few-Shot Visual Reasoning for Human-Object Interactions
http://openaccess.thecvf.com//content/CVPR2022/html/Jiang_Bongard-HOI_Benchmarking_Few-Shot_Visual_Reasoning_for_Human-Object_Interactions_CVPR_2022_paper.html
AUTHORS:		Huaizu Jiang, Xiaojian Ma, Weili Nie, Zhiding Yu, Yuke Zhu, Anima Anandkumar
HIGHLIGHT:		We introduce Bongard-HOI, a new visual reasoning benchmark that focuses on compositional learning of human-object interactions (HOIs) from natural images.

20, TITLE:		RIM-Net: Recursive Implicit Fields for Unsupervised Learning of Hierarchical Shape Structures
http://openaccess.thecvf.com//content/CVPR2022/html/Niu_RIM-Net_Recursive_Implicit_Fields_for_Unsupervised_Learning_of_Hierarchical_Shape_CVPR_2022_paper.html
AUTHORS:		Chengjie Niu, Manyi Li, Kai Xu, Hao Zhang
HIGHLIGHT:		We introduce RIM-Net, a neural network which learns recursive implicit fields for unsupervised inference of hierarchical shape structures.

21, TITLE: Do Learned Representations Respect Causal Relationships?
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Do_Learned_Representations_Respect_Causal_Relationships_CVPR_2022_paper.html
AUTHORS: Lan Wang, Vishnu Naresh Boddeti
HIGHLIGHT: Data often has many semantic attributes that are causally associated with each other. But do attribute-specific learned representations of data also respect the same causal relations? We answer this question in three steps.

22, TITLE: ZebraPose: Coarse To Fine Surface Encoding for 6DoF Object Pose Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Su_ZebraPose_Coarse_To_Fine_Surface_Encoding_for_6DoF_Object_Pose_CVPR_2022_paper.html
AUTHORS: Yongzhi Su, Mahdi Saleh, Torben Fetzer, Jason Rambach, Nassir Navab, Benjamin Busam, Didier Stricker, Federico Tombari
HIGHLIGHT: In this work, we present a discrete descriptor, which can represent the object surface densely.

23, TITLE: ZeroCap: Zero-Shot Image-to-Text Generation for Visual-Semantic Arithmetic
http://openaccess.thecvf.com//content/CVPR2022/html/Tewel_ZeroCap_Zero-Shot_Image-to-Text_Generation_for_Visual-Semantic_Arithmetic_CVPR_2022_paper.html
AUTHORS: Yoad Tewel, Yoav Shalev, Idan Schwartz, Lior Wolf
HIGHLIGHT: While such models can provide a powerful score for matching and subsequent zero-shot tasks, they are not capable of generating caption given an image. In this work, we repurpose such models to generate a descriptive text given an image at inference time, without any further training or tuning step.

24, TITLE: Learning To Affiliate: Mutual Centralized Learning for Few-Shot Classification
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Learning_To_Affiliate_Mutual_Centralized_Learning_for_Few-Shot_Classification_CVPR_2022_paper.html
AUTHORS: Yang Liu, Weifeng Zhang, Chao Xiang, Tu Zheng, Deng Cai, Xiaofei He
HIGHLIGHT: They generally explore a unidirectional paradigm, e.g., find the nearest support feature for every query feature and aggregate these local matches for a joint classification. In this paper, we propose a novel Mutual Centralized Learning (MCL) to fully affiliate these two disjoint dense features sets in a bidirectional paradigm.

25, TITLE: CAPRI-Net: Learning Compact CAD Shapes With Adaptive Primitive Assembly
http://openaccess.thecvf.com//content/CVPR2022/html/Yu_CAPRI-Net_Learning_Compact_CAD_Shapes_With_Adaptive_Primitive_Assembly_CVPR_2022_paper.html
AUTHORS: Fenggen Yu, Zhiqin Chen, Manyi Li, Aditya Sanghi, Hooman Shayani, Ali Mahdavi-Amiri, Hao Zhang
HIGHLIGHT: We introduce CAPRI-Net, a self-supervised neural network for learning compact and interpretable implicit representations of 3D computer-aided design (CAD) models, in the form of adaptive primitive assemblies.

26, TITLE: ATPFL: Automatic Trajectory Prediction Model Design Under Federated Learning Framework
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_ATPFL_Automatic_Trajectory_Prediction_Model_Design_Under_Federated_Learning_Framework_CVPR_2022_paper.html
AUTHORS: Chunnan Wang, Xiang Chen, Junzhe Wang, Hongzhi Wang
HIGHLIGHT: Besides, the existing works ignore Federated Learning (FL) scenarios, failing to make full use of distributed multi-source datasets with rich actual scenes to learn more a powerful TP model. In this paper, we make up for the above defects and propose ATPFL to help users federate multi-source trajectory datasets to automatically design and train a powerful TP model.

27, TITLE: Revisiting Learnable Affines for Batch Norm in Few-Shot Transfer Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Yazdanpanah_Revisiting_Learnable_Affines_for_Batch_Norm_in_Few-Shot_Transfer_Learning_CVPR_2022_paper.html
AUTHORS: Moslem Yazdanpanah, Aamer Abdul Rahman, Muawiz Chaudhary, Christian Desrosiers, Mohammad Havaei, Eugene Belilovsky, Samira Ebrahimi Kahou
HIGHLIGHT: These affine parameters were introduced to maintain the expressive powers of the model following normalization. While this hypothesis holds true for classification within the same domain, this work illustrates that these parameters are detrimental to downstream performance on common few-shot transfer tasks.

28, TITLE: Bridging the Gap Between Classification and Localization for Weakly Supervised Object Localization
http://openaccess.thecvf.com//content/CVPR2022/html/Kim_Bridging_the_Gap_Between_Classification_and_Localization_for_Weakly_Supervised_CVPR_2022_paper.html
AUTHORS: Eunji Kim, Siwon Kim, Jungbeom Lee, Hyunwoo Kim, Sungroh Yoon
HIGHLIGHT: In this work, we find the gap between classification and localization in terms of the misalignment of the directions between an input feature and a class-specific weight.

29, TITLE: Multi-Class Token Transformer for Weakly Supervised Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_Multi-Class_Token_Transformer_for_Weakly_Supervised_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Lian Xu, Wanli Ouyang, Mohammed Bennamoun, Farid Boussaid, Dan Xu
HIGHLIGHT: This paper proposes a new transformer-based framework to learn class-specific object localization maps as pseudo labels for weakly supervised semantic segmentation (WSSS).

30, TITLE: 3D Moments From Near-Duplicate Photos
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_3D_Moments_From_Near-Duplicate_Photos_CVPR_2022_paper.html

AUTHORS:		Qianqian Wang, Zhengqi Li, David Salesin, Noah Snavely, Brian Curless, Janne Kontkanen
HIGHLIGHT:		We introduce 3D Moments, a new computational photography effect.

31, TITLE:		Exact Feature Distribution Matching for Arbitrary Style Transfer and Domain Generalization
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Exact_Feature_Distribution_Matching_for_Arbitrary_Style_Transfer_and_Domain_CVPR_2022_paper.html
AUTHORS:		Yabin Zhang, Minghan Li, Ruihuang Li, Kui Jia, Lei Zhang
HIGHLIGHT:		In this work, we, for the first time to our best knowledge, propose to perform Exact Feature Distribution Matching (EFDM) by exactly matching the empirical Cumulative Distribution Functions (eCDFs) of image features, which could be implemented by applying the Exact Histogram Matching (EHM) in the image feature space.

32, TITLE:		Blind2Unblind: Self-Supervised Image Denoising With Visible Blind Spots
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Blind2Unblind_Self-Supervised_Image_Denoising_With_Visible_Blind_Spots_CVPR_2022_paper.html
AUTHORS:		Zejin Wang, Jiazheng Liu, Guoqing Li, Hua Han
HIGHLIGHT:		In this paper, we propose a simple yet efficient approach called Blind2Unblind to overcome the information loss in blindspot-driven denoising methods.

33, TITLE:		Balanced and Hierarchical Relation Learning for One-Shot Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Balanced_and_Hierarchical_Relation_Learning_for_One-Shot_Object_Detection_CVPR_2022_paper.html
AUTHORS:		Hanqing Yang, Sijia Cai, Hualian Sheng, Bing Deng, Jianqiang Huang, Xian-Sheng Hua, Yong Tang, Yu Zhang
HIGHLIGHT:		In this paper, we introduce the balanced and hierarchical learning for our detector.

34, TITLE:		End-to-End Generative Pretraining for Multimodal Video Captioning
http://openaccess.thecvf.com//content/CVPR2022/html/Seo_End-to-End_Generative_Pretraining_for_Multimodal_Video_Captioning_CVPR_2022_paper.html
AUTHORS:		Paul Hongsuck Seo, Arsha Nagrani, Anurag Arnab, Cordelia Schmid
HIGHLIGHT:		We present Multimodal Video Generative Pretraining (MV-GPT), a new pretraining framework for learning from unlabelled videos which can be effectively used for generative tasks such as multimodal video captioning.

35, TITLE:		Delving Deep Into the Generalization of Vision Transformers Under Distribution Shifts
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Delving_Deep_Into_the_Generalization_of_Vision_Transformers_Under_Distribution_CVPR_2022_paper.html
AUTHORS:		Chongzhi Zhang, Mingyuan Zhang, Shanghang Zhang, Daisheng Jin, Qiang Zhou, Zhongang Cai, Haiyu Zhao, Xianglong Liu, Ziwei Liu
HIGHLIGHT:		In this work, we provide a comprehensive study on the out-of-distribution generalization of Vision Transformers.

36, TITLE:		NICE-SLAM: Neural Implicit Scalable Encoding for SLAM
http://openaccess.thecvf.com//content/CVPR2022/html/Zhu_NICE-SLAM_Neural_Implicit_Scalable_Encoding_for_SLAM_CVPR_2022_paper.html
AUTHORS:		Zihan Zhu, Songyou Peng, Viktor Larsson, Weiwei Xu, Hujun Bao, Zhaopeng Cui, Martin R. Oswald, Marc Pollefeys
HIGHLIGHT:		In this paper, we present NICE-SLAM, a dense SLAM system that incorporates multi-level local information by introducing a hierarchical scene representation.

37, TITLE:		HyperDet3D: Learning a Scene-Conditioned 3D Object Detector
http://openaccess.thecvf.com//content/CVPR2022/html/Zheng_HyperDet3D_Learning_a_Scene-Conditioned_3D_Object_Detector_CVPR_2022_paper.html
AUTHORS:		Yu Zheng, Yueqi Duan, Jiwen Lu, Jie Zhou, Qi Tian
HIGHLIGHT:		In this paper, we propose HyperDet3D to explore scene-conditioned prior knowledge for 3D object detection.

38, TITLE:		Stochastic Trajectory Prediction via Motion Indeterminacy Diffusion
http://openaccess.thecvf.com//content/CVPR2022/html/Gu_Stochastic_Trajectory_Prediction_via_Motion_Indeterminacy_Diffusion_CVPR_2022_paper.html
AUTHORS:		Tianpei Gu, Guangyi Chen, Junlong Li, Chunze Lin, Yongming Rao, Jie Zhou, Jiwen Lu
HIGHLIGHT:		In this paper, we present a new framework to formulate the trajectory prediction task as a reverse process of motion indeterminacy diffusion (MID), in which we progressively discard indeterminacy from all the walkable areas until reaching the desired trajectory.

39, TITLE:		CLRNet: Cross Layer Refinement Network for Lane Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Zheng_CLRNet_Cross_Layer_Refinement_Network_for_Lane_Detection_CVPR_2022_paper.html
AUTHORS:		Tu Zheng, Yifei Huang, Yang Liu, Wenjian Tang, Zheng Yang, Deng Cai, Xiaofei He
HIGHLIGHT:		In this work, we present Cross Layer Refinement Network (CLRNet) aiming at fully utilizing both high-level and low-level features in lane detection.

40, TITLE:		Cross-Modal Map Learning for Vision and Language Navigation

http://openaccess.thecvf.com//content/CVPR2022/html/Georgakis_Cross-Modal_Map_Learning_for_Vision_and_Language_Navigation_CVPR_2022_paper.html
AUTHORS: Georgios Georgakis, Karl Schmeckpeper, Karan Wanchoo, Soham Dan, Eleni Miltsakaki, Dan Roth, Kostas Daniilidis
HIGHLIGHT: In this work, we propose a cross-modal map learning model for vision-and-language navigation that first learns to predict the top-down semantics on an egocentric map for both observed and unobserved regions, and then predicts a path towards the goal as a set of waypoints.

41, TITLE: Motion-Aware Contrastive Video Representation Learning via Foreground-Background Merging
http://openaccess.thecvf.com//content/CVPR2022/html/Ding_Motion-Aware_Contrastive_Video_Representation_Learning_via_Foreground-Background_Merging_CVPR_2022_paper.html
AUTHORS: Shuangrui Ding, Maomao Li, Tianyu Yang, Rui Qian, Haohang Xu, Qingyi Chen, Jue Wang, Hongkai Xiong
HIGHLIGHT: Such bias makes the model suffer from weak generalization ability, leading to worse performance on downstream tasks such as action recognition. To alleviate such bias, we propose Foreground-background Merging (FAME) to deliberately compose the moving foreground region of the selected video onto the static background of others.

42, TITLE: Incremental Transformer Structure Enhanced Image Inpainting With Masking Positional Encoding
http://openaccess.thecvf.com//content/CVPR2022/html/Dong_Incremental_Transformer_Structure_Enhanced_Image_Inpainting_With_Masking_Positional_Encoding_CVPR_2022_paper.html
AUTHORS: Qiaole Dong, Chenjie Cao, Yanwei Fu
HIGHLIGHT: On the other hand, attention-based models can learn better long-range dependency for the structure recovery, but they are limited by the heavy computation for inference with large image sizes. To address these issues, we propose to leverage an additional structure restorer to facilitate the image inpainting incrementally.

43, TITLE: Pointly-Supervised Instance Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Cheng_Pointly-Supervised_Instance_Segmentation_CVPR_2022_paper.html
AUTHORS: Bowen Cheng, Omkar Parkhi, Alexander Kirillov
HIGHLIGHT: We propose an embarrassingly simple point annotation scheme to collect weak supervision for instance segmentation.

44, TITLE: Cross-Modal Clinical Graph Transformer for Ophthalmic Report Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Cross-Modal_Clinical_Graph_Transformer_for_Ophthalmic_Report_Generation_CVPR_2022_paper.html
AUTHORS: Mingjie Li, Wenjia Cai, Karin Verspoor, Shirui Pan, Xiaodan Liang, Xiaojun Chang
HIGHLIGHT: To endow models with the capability of incorporating expert knowledge, we propose a Cross-modal clinical Graph Transformer (CGT) for ophthalmic report generation (ORG), in which clinical relation triples are injected into the visual features as prior knowledge to drive the decoding procedure.

45, TITLE: Human-Object Interaction Detection via Disentangled Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Zhou_Human-Object_Interaction_Detection_via_Disentangled_Transformer_CVPR_2022_paper.html
AUTHORS: Desen Zhou, Zhichao Liu, Jian Wang, Leshan Wang, Tao Hu, Errui Ding, Jingdong Wang
HIGHLIGHT: Our main motivation is that detecting the human-object instances and classifying interactions accurately needs to learn representations that focus on different regions. To this end, we present Disentangled Transformer, where both encoder and decoder are disentangled to facilitate learning of two subtasks.

46, TITLE: DINE: Domain Adaptation From Single and Multiple Black-Box Predictors
http://openaccess.thecvf.com//content/CVPR2022/html/Liang_DINE_Domain_Adaptation_From_Single_and_Multiple_Black-Box_Predictors_CVPR_2022_paper.html
AUTHORS: Jian Liang, Dapeng Hu, Jiashi Feng, Ran He
HIGHLIGHT: This paper studies a practical and interesting setting for UDA, where only black-box source models (i.e., only network predictions are available) are provided during adaptation in the target domain. To solve this problem, we propose a new two-step knowledge adaptation framework called DIstill and fine-tuNE (DINE).

47, TITLE: LGT-Net: Indoor Panoramic Room Layout Estimation With Geometry-Aware Transformer Network
http://openaccess.thecvf.com//content/CVPR2022/html/Jiang_LGT-Net_Indoor_Panoramic_Room_Layout_Estimation_With_Geometry-Aware_Transformer_Network_CVPR_2022_paper.html
AUTHORS: Zhigang Jiang, Zhongzheng Xiang, Jinhua Xu, Ming Zhao
HIGHLIGHT: We present that using horizon-depth along with room height can obtain omnidirectional-geometry awareness of room layout in both horizontal and vertical directions.

48, TITLE: CRIS: CLIP-Driven Referring Image Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_CRIS_CLIP-Driven_Referring_Image_Segmentation_CVPR_2022_paper.html
AUTHORS: Zhaoqing Wang, Yu Lu, Qiang Li, Xunqiang Tao, Yandong Guo, Mingming Gong, Tongliang Liu
HIGHLIGHT: Inspired by the recent advance in Contrastive Language-Image Pretraining (CLIP), in this paper, we propose an end-to-end CLIP-Driven Referring Image Segmentation framework (CRIS).

49, TITLE: Multi-View Mesh Reconstruction With Neural Deferred Shading

http://openaccess.thecvf.com//content/CVPR2022/html/Worchel_Multi-View_Mesh_Reconstruction_With_Neural_Deferred_Shading_CVPR_2022_paper.html
AUTHORS:		Markus Worchel, Rodrigo Diaz, Weiwen Hu, Oliver Schreer, Ingo Feldmann, Peter Eisert
HIGHLIGHT:		We propose an analysis-by-synthesis method for fast multi-view 3D reconstruction of opaque objects with arbitrary materials and illumination.

50, TITLE:		CVF-SID: Cyclic Multi-Variate Function for Self-Supervised Image Denoising by Disentangling Noise From Image
http://openaccess.thecvf.com//content/CVPR2022/html/Neshatavar_CVF-SID_Cyclic_Multi-Variate_Function_for_Self-Supervised_Image_Denoising_by_Disentangling_CVPR_2022_paper.html
AUTHORS:		Reyhaneh Neshatavar, Mohsen Yavartanoo, Sanghyun Son, Kyoung Mu Lee
HIGHLIGHT:		To address the aforementioned challenges, we propose a novel and powerful self-supervised denoising method called CVF-SID based on a Cyclic multi-Variate Function (CVF) module and a self-supervised image disentangling (SID) framework.

51, TITLE:		Infrared Invisible Clothing: Hiding From Infrared Detectors at Multiple Angles in Real World
http://openaccess.thecvf.com//content/CVPR2022/html/Zhu_Infrared_Invisible_Clothing_Hiding_From_Infrared_Detectors_at_Multiple_Angles_CVPR_2022_paper.html
AUTHORS:		Xiaopei Zhu, Zhanhao Hu, Siyuan Huang, Jianmin Li, Xiaolin Hu
HIGHLIGHT:		We proposed the infrared adversarial clothing, which could fool infrared pedestrian detectors at different angles.

52, TITLE:		Distribution-Aware Single-Stage Models for Multi-Person 3D Pose Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Distribution-Aware_Single-Stage_Models_for_Multi-Person_3D_Pose_Estimation_CVPR_2022_paper.html
AUTHORS:		Zitian Wang, Xuecheng Nie, Xiaochao Qu, Yunpeng Chen, Si Liu
HIGHLIGHT:		In this paper, we present a novel Distribution-Aware Single-stage (DAS) model for tackling the challenging multi-person 3D pose estimation problem.

53, TITLE:		FaceFormer: Speech-Driven 3D Facial Animation With Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Fan_FaceFormer_Speech-Driven_3D_Facial_Animation_With_Transformers_CVPR_2022_paper.html
AUTHORS:		Yingruo Fan, Zhaojiang Lin, Jun Saito, Wenping Wang, Taku Komura
HIGHLIGHT:		Prior works typically focus on learning phoneme-level features of short audio windows with limited context, occasionally resulting in inaccurate lip movements. To tackle this limitation, we propose a Transformer-based autoregressive model, FaceFormer, which encodes the long-term audio context and autoregressively predicts a sequence of animated 3D face meshes.

54, TITLE:		Exploring Patch-Wise Semantic Relation for Contrastive Learning in Image-to-Image Translation Tasks
http://openaccess.thecvf.com//content/CVPR2022/html/Jung_Exploring_Patch-Wise_Semantic_Relation_for_Contrastive_Learning_in_Image-to-Image_Translation_CVPR_2022_paper.html
AUTHORS:		Chanyong Jung, Gihyun Kwon, Jong Chul Ye
HIGHLIGHT:		However, the methods often ignore the diverse semantic relation within the images. To address this, here we propose a novel semantic relation consistency (SRC) regularization along with the decoupled contrastive learning (DCL), which utilize the diverse semantics by focusing on the heterogeneous semantics between the image patches of a single image.

55, TITLE:		High-Resolution Face Swapping via Latent Semantics Disentanglement
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_High-Resolution_Face_Swapping_via_Latent_Semantics_Disentanglement_CVPR_2022_paper.html
AUTHORS:		Yangyang Xu, Bailin Deng, Junle Wang, Yanqing Jing, Jia Pan, Shengfeng He
HIGHLIGHT:		We present a novel high-resolution face swapping method using the inherent prior knowledge of a pre-trained GAN model.

56, TITLE:		Searching the Deployable Convolution Neural Networks for GPUs
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Searching_the_Deployable_Convolution_Neural_Networks_for_GPUs_CVPR_2022_paper.html
AUTHORS:		Linnan Wang, Chenhan Yu, Satish Salian, Slawomir Kierat, Szymon Migacz, Alex Fit Florea
HIGHLIGHT:		This paper intends to expedite the model customization with a model hub that contains the optimized models tiered by their inference latency using Neural Architecture Search (NAS).

57, TITLE:		Sparse Local Patch Transformer for Robust Face Alignment and Landmarks Inherent Relation Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Xia_Sparse_Local_Patch_Transformer_for_Robust_Face_Alignment_and_Landmarks_CVPR_2022_paper.html
AUTHORS:		Jiahao Xia, Weiwei Qu, Wenjian Huang, Jianguo Zhang, Xi Wang, Min Xu
HIGHLIGHT:		In this paper, we propose a Sparse Local Patch Transformer (SLPT) for learning the inherent relation.

58, TITLE:		DeepFake Disrupter: The Detector of DeepFake Is My Friend
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_DeepFake_Disrupter_The_Detector_of_DeepFake_Is_My_Friend_CVPR_2022_paper.html
AUTHORS:		Xueyu Wang, Jiajun Huang, Siqi Ma, Surya Nepal, Chang Xu
HIGHLIGHT:		In this paper, we propose a novel DeepFake disruption algorithm called "DeepFake Disrupter".

59, TITLE:		Rotationally Equivariant 3D Object Detection

http://openaccess.thecvf.com//content/CVPR2022/html/Yu_Rotationally_Equivariant_3D_Object_Detection_CVPR_2022_paper.html
AUTHORS: Hong-Xing Yu, Jiajun Wu, Li Yi
HIGHLIGHT: To incorporate object-level rotation equivariance into 3D object detectors, we need a mechanism to extract equivariant features with local object-level spatial support while being able to model cross-object context information. To this end, we propose Equivariant Object detection Network (EON) with a rotation equivariance suspension design to achieve object-level equivariance.

60, TITLE: Accelerating DETR Convergence via Semantic-Aligned Matching
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Accelerating_DETR_Convergence_via_Semantic-Aligned_Matching_CVPR_2022_paper.html
AUTHORS: Gongjie Zhang, Zhipeng Luo, Yingchen Yu, Kaiwen Cui, Shijian Lu
HIGHLIGHT: We observe that the slow convergence is largely attributed to the complication in matching object queries with target features in different feature embedding spaces. This paper presents SAM-DETR, a Semantic-Aligned-Matching DETR that greatly accelerates DETR's convergence without sacrificing its accuracy.

61, TITLE: Long-Short Temporal Contrastive Learning of Video Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Long-Short_Temporal_Contrastive_Learning_of_Video_Transformers_CVPR_2022_paper.html
AUTHORS: Jue Wang, Gedas Bertasius, Du Tran, Lorenzo Torresani
HIGHLIGHT: In this paper, we empirically demonstrate that self-supervised pretraining of video transformers on video-only datasets can lead to action recognition results that are on par or better than those obtained with supervised pretraining on large-scale image datasets, even massive ones such as ImageNet-21K.

62, TITLE: Vision Transformer With Deformable Attention
http://openaccess.thecvf.com//content/CVPR2022/html/Xia_Vision_Transformer_With_Deformable_Attention_CVPR_2022_paper.html
AUTHORS: Zhuofan Xia, Xuran Pan, Shiji Song, Li Erran Li, Gao Huang
HIGHLIGHT: This flexible scheme enables the self-attention module to focus on relevant regions and cap-ture more informative features. On this basis, we present Deformable Attention Transformer, a general backbone model with deformable attention for both image classifi-cation and dense prediction tasks.

63, TITLE: Towards General Purpose Vision Systems: An End-to-End Task-Agnostic Vision-Language Architecture
http://openaccess.thecvf.com//content/CVPR2022/html/Gupta_Towards_General_Purpose_Vision_Systems_An_End-to-End_Task-Agnostic_Vision-Language_Architecture_CVPR_2022_paper.html
AUTHORS: Tanmay Gupta, Amita Kamath, Aniruddha Kembhavi, Derek Hoiem
HIGHLIGHT: In this paper, we propose GPV-1, a task-agnostic vision-language architecture that can learn and perform tasks that involve receiving an image and producing text and/or bounding boxes, including classification, localization, visual question answering, captioning, and more.

64, TITLE: Deep Vanishing Point Detection: Geometric Priors Make Dataset Variations Vanish
http://openaccess.thecvf.com//content/CVPR2022/html/Lin_Deep_Vanishing_Point_Detection_Geometric_Priors_Make_Dataset_Variations_Vanish_CVPR_2022_paper.html
AUTHORS: Yancong Lin, Ruben Wiersma, Silvia L. Pintea, Klaus Hildebrandt, Elmar Eisemann, Jan C. van Gemert
HIGHLIGHT: Yet, deep networks require expensive annotated datasets trained on costly hardware and do not generalize to even slightly different domains, and minor problem variants. Here, we address these issues by injecting deep vanishing point detection networks with prior knowledge.

65, TITLE: RM-Depth: Unsupervised Learning of Recurrent Monocular Depth in Dynamic Scenes
http://openaccess.thecvf.com//content/CVPR2022/html/Hui_RM-Depth_Unsupervised_Learning_of_Recurrent_Monocular_Depth_in_Dynamic_Scenes_CVPR_2022_paper.html
AUTHORS: Tak-Wai Hui
HIGHLIGHT: In this paper, an unsupervised learning framework is proposed to jointly predict monocular depth and complete 3D motion including the motions of moving objects and camera.

66, TITLE: LiT: Zero-Shot Transfer With Locked-Image Text Tuning
http://openaccess.thecvf.com//content/CVPR2022/html/Zhai_LiT_Zero-Shot_Transfer_With_Locked-Image_Text_Tuning_CVPR_2022_paper.html
AUTHORS: Xiaohua Zhai, Xiao Wang, Basil Mustafa, Andreas Steiner, Daniel Keysers, Alexander Kolesnikov, Lucas Beyer
HIGHLIGHT: This paper presents contrastive-tuning, a simple method employing contrastive training to align image and text models while still taking advantage of their pre-training.

67, TITLE: Cloning Outfits From Real-World Images to 3D Characters for Generalizable Person Re-Identification
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Cloning_Outfits_From_Real-World_Images_to_3D_Characters_for_Generalizable_CVPR_2022_paper.html
AUTHORS: Yanan Wang, Xuezhi Liang, Shengcai Liao
HIGHLIGHT: However, synthesized persons in existing datasets are mostly cartoon-like and in random dress collocation, which limits their performance. To address this, in this work, an automatic approach is proposed to directly clone the whole outfits from real-world person images to virtual 3D characters, such that any virtual person thus created will appear very similar to its real-world counterpart.

68, TITLE: GeoNeRF: Generalizing NeRF With Geometry Priors
http://openaccess.thecvf.com//content/CVPR2022/html/Johari_GeoNeRF_Generalizing_NeRF_With_Geometry_Priors_CVPR_2022_paper.html
AUTHORS: Mohammad Mahdi Johari, Yann Lepoittevin, François Fleuret
HIGHLIGHT: We present GeoNeRF, a generalizable photorealistic novel view synthesis method based on neural radiance fields.

69, TITLE: ABPN: Adaptive Blend Pyramid Network for Real-Time Local Retouching of Ultra High-Resolution Photo
http://openaccess.thecvf.com//content/CVPR2022/html/Lei_ABPN_Adaptive_Blend_Pyramid_Network_for_Real-Time_Local_Retouching_of_CVPR_2022_paper.html
AUTHORS: Biwen Lei, Xiefan Guo, Hongyu Yang, Miaomiao Cui, Xuansong Xie, Di Huang
HIGHLIGHT: In this paper, we propose a novel adaptive blend pyramid network, which aims to achieve fast local retouching on ultra high-resolution photos.

70, TITLE: PhoCaL: A Multi-Modal Dataset for Category-Level Object Pose Estimation With Photometrically Challenging Objects
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_PhoCaL_A_Multi-Modal_Dataset_for_Category-Level_Object_Pose_Estimation_With_CVPR_2022_paper.html
AUTHORS: Pengyuan Wang, HyunJun Jung, Yitong Li, Siyuan Shen, Rahul Parthasarathy Srikanth, Lorenzo Garattoni, Sven Meier, Nassir Navab, Benjamin Busam
HIGHLIGHT: To provide a benchmark with high-quality ground truth annotations to the community, we introduce a multimodal dataset for category-level object pose estimation with photometrically challenging objects termed PhoCaL.

71, TITLE: Neural Compression-Based Feature Learning for Video Restoration
http://openaccess.thecvf.com//content/CVPR2022/html/Huang_Neural_Compression-Based_Feature_Learning_for_Video_Restoration_CVPR_2022_paper.html
AUTHORS: Cong Huang, Jiahao Li, Bin Li, Dong Liu, Yan Lu
HIGHLIGHT: This paper proposes learning noise-robust feature representations to help video restoration.

72, TITLE: Expanding Low-Density Latent Regions for Open-Set Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Han_Expanding_Low-Density_Latent_Regions_for_Open-Set_Object_Detection_CVPR_2022_paper.html
AUTHORS: Jiaming Han, Yuqiang Ren, Jian Ding, Xingjia Pan, Ke Yan, Gui-Song Xia
HIGHLIGHT: In this work, we propose to identify unknown objects by separating high/low-density regions in the latent space, based on the consensus that unknown objects are usually distributed in low-density latent regions.

73, TITLE: Drop the GAN: In Defense of Patches Nearest Neighbors As Single Image Generative Models
http://openaccess.thecvf.com//content/CVPR2022/html/Granot_Drop_the_GAN_In_Defense_of_Patches_Nearest_Neighbors_As_CVPR_2022_paper.html
AUTHORS: Niv Granot, Ben Feinstein, Assaf Shocher, Shai Bagon, Michal Irani
HIGHLIGHT: However, despite their impressiveness, single-image GANs require long training time (usually hours) for each image and each task and often suffer from visual artifacts. In this paper we revisit the classical patch-based methods, and show that - unlike previously believed -- classical methods can be adapted to tackle these novel "GAN-only" tasks.

74, TITLE: Uformer: A General U-Shaped Transformer for Image Restoration
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Uformer_A_General_U-Shaped_Transformer_for_Image_Restoration_CVPR_2022_paper.html
AUTHORS: Zhendong Wang, Xiaodong Cun, Jianmin Bao, Wengang Zhou, Jianzhuang Liu, Houqiang Li
HIGHLIGHT: In this paper, we present Uformer, an effective and efficient Transformer-based architecture for image restoration, in which we build a hierarchical encoder-decoder network using the Transformer block.

75, TITLE: Exploring Dual-Task Correlation for Pose Guided Person Image Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Exploring_Dual-Task_Correlation_for_Pose_Guided_Person_Image_Generation_CVPR_2022_paper.html
AUTHORS: Pengze Zhang, Lingxiao Yang, Jian-Huang Lai, Xiaohua Xie
HIGHLIGHT: Most of the existing methods only focus on the ill-posed source-to-target task and fail to capture reasonable texture mapping. To address this problem, we propose a novel Dual-task Pose Transformer Network (DPTN), which introduces an auxiliary task (i.e., source-tosource task) and exploits the dual-task correlation to promote the performance of PGPIG.

76, TITLE: Portrait Eyeglasses and Shadow Removal by Leveraging 3D Synthetic Data
http://openaccess.thecvf.com//content/CVPR2022/html/Lyu_Portrait_Eyeglasses_and_Shadow_Removal_by_Leveraging_3D_Synthetic_Data_CVPR_2022_paper.html
AUTHORS: Junfeng Lyu, Zhibo Wang, Feng Xu
HIGHLIGHT: In this paper, we propose a novel framework to remove eyeglasses as well as their cast shadows from face images.

77, TITLE: Neural Rays for Occlusion-Aware Image-Based Rendering
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Neural_Rays_for_Occlusion-Aware_Image-Based_Rendering_CVPR_2022_paper.html

AUTHORS: Yuan Liu, Sida Peng, Lingjie Liu, Qianqian Wang, Peng Wang, Christian Theobalt, Xiaowei Zhou, Wenping Wang

HIGHLIGHT: We present a new neural representation, called Neural Ray (NeuRay), for the novel view synthesis task.

78, TITLE: Modeling 3D Layout for Group Re-Identification
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Modeling_3D_Layout_for_Group_Re-Identification_CVPR_2022_paper.html
AUTHORS: Quan Zhang, Kaiheng Dang, Jian-Huang Lai, Zhanxiang Feng, Xiaohua Xie
HIGHLIGHT: However, layout ambiguity is introduced because these methods only consider the 2D layout on the imaging plane. In this paper, we overcome the above limitations by 3D layout modeling.

79, TITLE: Open-World Instance Segmentation: Exploiting Pseudo Ground Truth From Learned Pairwise Affinity
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Open-World_Instance_Segmentation_Exploiting_Pseudo_Ground_Truth_From_Learned_Pairwise_CVPR_2022_paper.html
AUTHORS: Weiyao Wang, Matt Feiszli, Heng Wang, Jitendra Malik, Du Tran
HIGHLIGHT: Here we propose a novel approach for mask proposals, Generic Grouping Networks (GGNs), constructed without semantic supervision.

80, TITLE: SIOD: Single Instance Annotated per Category per Image for Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Li_SIOD_Single_Instance_Annotated_per_Category_per_Image_for_Object_CVPR_2022_paper.html
AUTHORS: Hanjun Li, Xingjia Pan, Ke Yan, Fan Tang, Wei-Shi Zheng
HIGHLIGHT: Under the SIOD setting, we propose a simple yet effective framework, termed Dual-Mining (DMiner), which consists of a Similarity-based Pseudo Label Generating module (SPLG) and a Pixel-level Group Contrastive Learning module (PGCL).

81, TITLE: Toward Fast, Flexible, and Robust Low-Light Image Enhancement
http://openaccess.thecvf.com//content/CVPR2022/html/Ma_Toward_Fast_Flexible_and_Robust_Low-Light_Image_Enhancement_CVPR_2022_paper.html
AUTHORS: Long Ma, Tengyu Ma, Risheng Liu, Xin Fan, Zhongxuan Luo
HIGHLIGHT: In this paper, we develop a new Self-Calibrated Illumination (SCI) learning framework for fast, flexible, and robust brightening images in real-world low-light scenarios.

82, TITLE: Online Learning of Reusable Abstract Models for Object Goal Navigation
http://openaccess.thecvf.com//content/CVPR2022/html/Campari_Online_Learning_of_Reusable_Abstract_Models_for_Object_Goal_Navigation_CVPR_2022_paper.html
AUTHORS: Tommaso Campari, Leonardo Lamanna, Paolo Traverso, Luciano Serafini, Lamberto Ballan
HIGHLIGHT: In this paper, we present a novel approach to incrementally learn an Abstract Model of an unknown environment, and show how an agent can reuse the learned model for tackling the Object Goal Navigation task.

83, TITLE: Bridge-Prompt: Towards Ordinal Action Understanding in Instructional Videos
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Bridge-Prompt_Towards_Ordinal_Action_Understanding_in_Instructional_Videos_CVPR_2022_paper.html
AUTHORS: Muheng Li, Lei Chen, Yueqi Duan, Zhilan Hu, Jianjiang Feng, Jie Zhou, Jiwen Lu
HIGHLIGHT: In this paper, we propose a prompt-based framework, Bridge-Prompt (Br-Prompt), to model the semantics across adjacent actions, so that it simultaneously exploits both out-of-context and contextual information from a series of ordinal actions in instructional videos.

84, TITLE: SimMatch: Semi-Supervised Learning With Similarity Matching
http://openaccess.thecvf.com//content/CVPR2022/html/Zheng_SimMatch_Semi-Supervised_Learning_With_Similarity_Matching_CVPR_2022_paper.html
AUTHORS: Mingkai Zheng, Shan You, Lang Huang, Fei Wang, Chen Qian, Chang Xu
HIGHLIGHT: In this paper, we introduced a new semi-supervised learning framework, SimMatch, which simultaneously considers semantic similarity and instance similarity.

85, TITLE: OrphicX: A Causality-Inspired Latent Variable Model for Interpreting Graph Neural Networks
http://openaccess.thecvf.com//content/CVPR2022/html/Lin_OrphicX_A_Causality-Inspired_Latent_Variable_Model_for_Interpreting_Graph_Neural_CVPR_2022_paper.html
AUTHORS: Wanyu Lin, Hao Lan, Hao Wang, Baochun Li
HIGHLIGHT: This paper proposes a new eXplanation framework, called OrphicX, for generating causal explanations for any graph neural networks (GNNs) based on learned latent causal factors.

86, TITLE: HandOccNet: Occlusion-Robust 3D Hand Mesh Estimation Network
http://openaccess.thecvf.com//content/CVPR2022/html/Park_HandOccNet_Occlusion-Robust_3D_Hand_Mesh_Estimation_Network_CVPR_2022_paper.html
AUTHORS: JoonKyu Park, Yeonguk Oh, Gyeongsik Moon, Hongsuk Choi, Kyoung Mu Lee
HIGHLIGHT: Thus, in this work, we propose a novel 3D hand mesh estimation network HandOccNet, that can fully exploits the information at occluded regions as a secondary means to enhance image features and make it much richer.

87, TITLE: EfficientNeRF Efficient Neural Radiance Fields

http://openaccess.thecvf.com//content/CVPR2022/html/Hu_EfficientNeRF__Efficient_Neural_Radiance_Fields_CVPR_2022_paper.html
AUTHORS: Tao Hu, Shu Liu, Yilun Chen, Tiancheng Shen, Jiaya Jia
HIGHLIGHT: In this paper, we present EfficientNeRF as an efficient NeRF-based method to represent 3D scene and synthesize novel-view images.

88, TITLE: Quantifying Societal Bias Amplification in Image Captioning
http://openaccess.thecvf.com//content/CVPR2022/html/Hirota_Quantifying_Societal_Bias_Amplification_in_Image_Captioning_CVPR_2022_paper.html
AUTHORS: Yusuke Hirota, Yuta Nakashima, Noa Garcia
HIGHLIGHT: We provide a comprehensive study on the strengths and limitations of each metric, and propose LIC, a metric to study captioning bias amplification.

89, TITLE: Modular Action Concept Grounding in Semantic Video Prediction
http://openaccess.thecvf.com//content/CVPR2022/html/Yu_Modular_Action_Concept_Grounding_in_Semantic_Video_Prediction_CVPR_2022_paper.html
AUTHORS: Wei Yu, Wenxin Chen, Songheng Yin, Steve Easterbrook, Animesh Garg
HIGHLIGHT: Inspired by the idea of Mixture of Experts, we embody each abstract label by a structured combination of various visual concept learners and propose a novel video prediction model, Modular Action Concept Network (MAC).

90, TITLE: StyleSwin: Transformer-Based GAN for High-Resolution Image Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_StyleSwin_Transformer-Based_GAN_for_High-Resolution_Image_Generation_CVPR_2022_paper.html
AUTHORS: Bowen Zhang, Shuyang Gu, Bo Zhang, Jianmin Bao, Dong Chen, Fang Wen, Yong Wang, Baining Guo
HIGHLIGHT: In this paper, we seek to explore using pure transformers to build a generative adversarial network for high-resolution image synthesis.

91, TITLE: Reinforced Structured State-Evolution for Vision-Language Navigation
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Reinforced_Structured_State-Evolution_for_Vision-Language_Navigation_CVPR_2022_paper.html
AUTHORS: Jinyu Chen, Chen Gao, Erli Meng, Qiong Zhang, Si Liu
HIGHLIGHT: In this paper, we propose a novel Structured state-Evolution (SEvol) model to effectively maintain the environment layout clues for VLN.

92, TITLE: Sub-Word Level Lip Reading With Visual Attention
http://openaccess.thecvf.com//content/CVPR2022/html/Prajwal_Sub-Word_Level_Lip_Reading_With_Visual_Attention_CVPR_2022_paper.html
AUTHORS: K R Prajwal, Triantafyllos Afouras, Andrew Zisserman
HIGHLIGHT: The goal of this paper is to learn strong lip reading models that can recognise speech in silent videos.

93, TITLE: Weakly Supervised High-Fidelity Clothing Model Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Feng_Weakly_Supervised_High-Fidelity_Clothing_Model_Generation_CVPR_2022_paper.html
AUTHORS: Ruili Feng, Cheng Ma, Chengji Shen, Xin Gao, Zhenjiang Liu, Xiaobo Li, Kairi Ou, Deli Zhao, Zheng-Jun Zha
HIGHLIGHT: However, the expensive proprietary model images challenge the existing image virtual try-on methods in this scenario, as most of them need to be trained on considerable amounts of model images accompanied with paired clothes images. In this paper, we propose a cheap yet scalable weakly-supervised method called Deep Generative Projection (DGP) to address this specific scenario.

94, TITLE: Highly-Efficient Incomplete Large-Scale Multi-View Clustering With Consensus Bipartite Graph
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Highly-Efficient_Incomplete_Large-Scale_Multi-View_Clustering_With_Consensus_Bipartite_Graph_CVPR_2022_paper.html
AUTHORS: Siwei Wang, Xinwang Liu, Li Liu, Wenxuan Tu, Xinzhong Zhu, Jiyuan Liu, Sihang Zhou, En Zhu
HIGHLIGHT: Although many IMVC methods have been recently proposed, they always encounter high complexity and expensive time expenditure from being applied into large-scale tasks. In this paper, we present a flexible highly-efficient incomplete large-scale multi-view clustering approach based on bipartite graph framework to solve these issues.

95, TITLE: Towards Principled Disentanglement for Domain Generalization
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Towards_Principled_Disentanglement_for_Domain_Generalization_CVPR_2022_paper.html
AUTHORS: Hanlin Zhang, Yi-Fan Zhang, Weiyang Liu, Adrian Weller, Bernhard Schölkopf, Eric P. Xing
HIGHLIGHT: Based on the transformation, we propose a primal-dual algorithm for joint representation disentanglement and domain generalization.

96, TITLE: Discrete Cosine Transform Network for Guided Depth Map Super-Resolution
http://openaccess.thecvf.com//content/CVPR2022/html/Zhao_Discrete_Cosine_Transform_Network_for_Guided_Depth_Map_Super-Resolution_CVPR_2022_paper.html
AUTHORS: Zixiang Zhao, Jiangshe Zhang, Shuang Xu, Zudi Lin, Hanspeter Pfister

HIGHLIGHT:　　　　To solve the challenges in interpreting the working mechanism, extracting cross-modal features and RGB texture over-transferred, we propose a novel Discrete Cosine Transform Network (DCTNet) to alleviate the problems from three aspects.

97, TITLE:　　　　Cerberus Transformer: Joint Semantic, Affordance and Attribute Parsing
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Cerberus_Transformer_Joint_Semantic_Affordance_and_Attribute_Parsing_CVPR_2022_paper.html
AUTHORS:　　　　Xiaoxue Chen, Tianyu Liu, Hao Zhao, Guyue Zhou, Ya-Qin Zhang
HIGHLIGHT:　　　　In this paper, we tackle the new problem of joint semantic, affordance and attribute parsing.

98, TITLE:　　　　E2V-SDE: From Asynchronous Events to Fast and Continuous Video Reconstruction via Neural Stochastic Differential Equations
http://openaccess.thecvf.com//content/CVPR2022/html/Kim_E2V-SDE_From_Asynchronous_Events_to_Fast_and_Continuous_Video_Reconstruction_CVPR_2022_paper.html
AUTHORS:　　　　Jongwan Kim, DongJin Lee, Byunggook Na, Seongsik Park, Sungroh Yoon
HIGHLIGHT:　　　　However, these works do not provide videos with sufficiently good quality due to unrealistic artifacts, such as lack of temporal information from irregular and discontinuous data and deterministic modeling for continuous-time stochastic process. In this study, we overcome these difficulties by introducing a new model called E2V-SDE, which is a neural continuous time-state model consisting of a latent stochastic differential equation and a conditional distribution of the observation.

99, TITLE:　　　　CoSSL: Co-Learning of Representation and Classifier for Imbalanced Semi-Supervised Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Fan_CoSSL_Co-Learning_of_Representation_and_Classifier_for_Imbalanced_Semi-Supervised_Learning_CVPR_2022_paper.html
AUTHORS:　　　　Yue Fan, Dengxin Dai, Anna Kukleva, Bernt Schiele
HIGHLIGHT:　　　　However, the more realistic setting of class-imbalanced data - called imbalanced SSL - is largely underexplored and standard SSL tends to underperform. In this paper, we propose a novel co-learning framework (CoSSL), which decouples representation and classifier learning while coupling them closely.

100, TITLE:　　　　Discovering Objects That Can Move
http://openaccess.thecvf.com//content/CVPR2022/html/Bao_Discovering_Objects_That_Can_Move_CVPR_2022_paper.html
AUTHORS:　　　　Zhipeng Bao, Pavel Tokmakov, Allan Jabri, Yu-Xiong Wang, Adrien Gaidon, Martial Hebert
HIGHLIGHT:　　　　This paper studies the problem of object discovery -- separating objects from the background without manual labels.

101, TITLE:　　　　Knowledge Mining With Scene Text for Fine-Grained Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Knowledge_Mining_With_Scene_Text_for_Fine-Grained_Recognition_CVPR_2022_paper.html
AUTHORS:　　　　Hao Wang, Junchao Liao, Tianheng Cheng, Zewen Gao, Hao Liu, Bo Ren, Xiang Bai, Wenyu Liu
HIGHLIGHT:　　　　We propose an end-to-end trainable network that mines implicit contextual knowledge behind scene text image and enhance the semantics and correlation to fine-tune the image representation.

102, TITLE:　　　　Self-Supervised Learning of Object Parts for Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Ziegler_Self-Supervised_Learning_of_Object_Parts_for_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS:　　　　Adrian Ziegler, Yuki M. Asano
HIGHLIGHT:　　　　However, learning dense representations is challenging, as in the unsupervised context it is not clear how to guide the model to learn representations that correspond to various potential object categories. In this paper, we argue that self-supervised learning of object parts is a solution to this issue.

103, TITLE:　　　　Iterative Corresponding Geometry: Fusing Region and Depth for Highly Efficient 3D Tracking of Textureless Objects
http://openaccess.thecvf.com//content/CVPR2022/html/Stoiber_Iterative_Corresponding_Geometry_Fusing_Region_and_Depth_for_Highly_Efficient_CVPR_2022_paper.html
AUTHORS:　　　　Manuel Stoiber, Martin Sundermeyer, Rudolph Triebel
HIGHLIGHT:　　　　In the following, we thus propose ICG, a novel probabilistic tracker that fuses region and depth information and only requires the object geometry.

104, TITLE:　　　　Single-Photon Structured Light
http://openaccess.thecvf.com//content/CVPR2022/html/Sundar_Single-Photon_Structured_Light_CVPR_2022_paper.html
AUTHORS:　　　　Varun Sundar, Sizhuo Ma, Aswin C. Sankaranarayanan, Mohit Gupta
HIGHLIGHT:　　　　We present a novel structured light technique that uses Single Photon Avalanche Diode (SPAD) arrays to enable 3D scanning at high-frame rates and low-light levels.

105, TITLE:　　　　Deblurring via Stochastic Refinement
http://openaccess.thecvf.com//content/CVPR2022/html/Whang_Deblurring_via_Stochastic_Refinement_CVPR_2022_paper.html
AUTHORS:　　　　Jay Whang, Mauricio Delbracio, Hossein Talebi, Chitwan Saharia, Alexandros G. Dimakis, Peyman Milanfar
HIGHLIGHT:　　　　We present an alternative framework for blind deblurring based on conditional diffusion models.

106, TITLE:　　　　3DJCG: A Unified Framework for Joint Dense Captioning and Visual Grounding on 3D Point Clouds

http://openaccess.thecvf.com//content/CVPR2022/html/Cai_3DJCG_A_Unified_Framework_for_Joint_Dense_Captioning_and_Visual_CVPR_2022_paper.html
AUTHORS: Daigang Cai, Lichen Zhao, Jing Zhang, Lu Sheng, Dong Xu
HIGHLIGHT: Observing that the 3D captioning task and the 3D grounding task contain both shared and complementary information in nature, in this work, we propose a unified framework to jointly solve these two distinct but closely related tasks in a synergistic fashion, which consists of both shared task-agnostic modules and lightweight task-specific modules.

107, TITLE: TransGeo: Transformer Is All You Need for Cross-View Image Geo-Localization
http://openaccess.thecvf.com//content/CVPR2022/html/Zhu_TransGeo_Transformer_Is_All_You_Need_for_Cross-View_Image_Geo-Localization_CVPR_2022_paper.html
AUTHORS: Sijie Zhu, Mubarak Shah, Chen Chen
HIGHLIGHT: The dominant CNN-based methods for cross-view image geo-localization rely on polar transform and fail to model global correlation. We propose a pure transformer-based approach (TransGeo) to address these limitations from a different perspective.

108, TITLE: R(Det)2: Randomized Decision Routing for Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Li_RDet2_Randomized_Decision_Routing_for_Object_Detection_CVPR_2022_paper.html
AUTHORS: Yali Li, Shengjin Wang
HIGHLIGHT: In this paper, we propose a novel approach to combine decision trees and deep neural networks in an end-to-end learning manner for object detection.

109, TITLE: Abandoning the Bayer-Filter To See in the Dark
http://openaccess.thecvf.com//content/CVPR2022/html/Dong_Abandoning_the_Bayer-Filter_To_See_in_the_Dark_CVPR_2022_paper.html
AUTHORS: Xingbo Dong, Wanyan Xu, Zhihui Miao, Lan Ma, Chao Zhang, Jiewen Yang, Zhe Jin, Andrew Beng Jin Teoh, Jiajun Shen
HIGHLIGHT: Due to the fact that not all photons can pass the Bayer-Filter on the sensor of the color camera, in this work, we first present a De-Bayer-Filter simulator based on deep neural networks to generate a monochrome raw image from the colored raw image. Next, a fully convolutional network is proposed to achieve the low-light image enhancement by fusing colored raw data with synthesized monochrome data.

110, TITLE: SASIC: Stereo Image Compression With Latent Shifts and Stereo Attention
http://openaccess.thecvf.com//content/CVPR2022/html/Wodlinger_SASIC_Stereo_Image_Compression_With_Latent_Shifts_and_Stereo_Attention_CVPR_2022_paper.html
AUTHORS: Matthias Wödlinger, Jan Kotera, Jan Xu, Robert Sablatnig
HIGHLIGHT: We propose a learned method for stereo image compression that leverages the similarity of the left and right images in a stereo pair due to overlapping fields of view.

111, TITLE: Exploiting Temporal Relations on Radar Perception for Autonomous Driving
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Exploiting_Temporal_Relations_on_Radar_Perception_for_Autonomous_Driving_CVPR_2022_paper.html
AUTHORS: Peizhao Li, Pu Wang, Karl Berntorp, Hongfu Liu
HIGHLIGHT: To enhance the capacity of automotive radar, in this work, we exploit the temporal information from successive ego-centric bird-eye-view radar image frames for radar object recognition.

112, TITLE: Multi-Instance Point Cloud Registration by Efficient Correspondence Clustering
http://openaccess.thecvf.com//content/CVPR2022/html/Tang_Multi-Instance_Point_Cloud_Registration_by_Efficient_Correspondence_Clustering_CVPR_2022_paper.html
AUTHORS: Weixuan Tang, Danping Zou
HIGHLIGHT: We propose to directly group the set of noisy correspondences into different clusters based on a distance invariance matrix.

113, TITLE: Contrastive Boundary Learning for Point Cloud Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Tang_Contrastive_Boundary_Learning_for_Point_Cloud_Segmentation_CVPR_2022_paper.html
AUTHORS: Liyao Tang, Yibing Zhan, Zhe Chen, Baosheng Yu, Dacheng Tao
HIGHLIGHT: In this paper, we focus on the segmentation of scene boundaries.

114, TITLE: Details or Artifacts: A Locally Discriminative Learning Approach to Realistic Image Super-Resolution
http://openaccess.thecvf.com//content/CVPR2022/html/Liang_Details_or_Artifacts_A_Locally_Discriminative_Learning_Approach_to_Realistic_CVPR_2022_paper.html
AUTHORS: Jie Liang, Hui Zeng, Lei Zhang
HIGHLIGHT: In this paper, we demonstrate that it is possible to train a GAN-based SISR model which can stably generate perceptually realistic details while inhibiting visual artifacts.

115, TITLE: CVNet: Contour Vibration Network for Building Extraction
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_CVNet_Contour_Vibration_Network_for_Building_Extraction_CVPR_2022_paper.html
AUTHORS: Ziqiang Xu, Chunyan Xu, Zhen Cui, Xiangwei Zheng, Jian Yang

HIGHLIGHT: Inspired by the physical vibration theory, we propose a contour vibration network (CVNet) for automatic building boundary delineation.

116, TITLE: Hyperbolic Image Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Atigh_Hyperbolic_Image_Segmentation_CVPR_2022_paper.html
AUTHORS: Mina Ghadimi Atigh, Julian Schoep, Erman Acar, Nanne van Noord, Pascal Mettes
HIGHLIGHT: In this work, we show that hyperbolic manifolds provide a valuable alternative for image segmentation and propose a tractable formulation of hierarchical pixel-level classification in hyperbolic space.

117, TITLE: Forward Compatible Training for Large-Scale Embedding Retrieval Systems
http://openaccess.thecvf.com//content/CVPR2022/html/Ramanujan_Forward_Compatible_Training_for_Large-Scale_Embedding_Retrieval_Systems_CVPR_2022_paper.html
AUTHORS: Vivek Ramanujan, Pavan Kumar Anasosalu Vasu, Ali Farhadi, Oncel Tuzel, Hadi Pouransari
HIGHLIGHT: In this work, we propose a new learning paradigm for representation learning: forward compatible training (FCT).

118, TITLE: Everything at Once - Multi-Modal Fusion Transformer for Video Retrieval
http://openaccess.thecvf.com//content/CVPR2022/html/Shvetsova_Everything_at_Once_-_Multi-Modal_Fusion_Transformer_for_Video_Retrieval_CVPR_2022_paper.html
AUTHORS: Nina Shvetsova, Brian Chen, Andrew Rouditchenko, Samuel Thomas, Brian Kingsbury, Rogerio S. Feris, David Harwath, James Glass, Hilde Kuehne
HIGHLIGHT: Multi-modal learning from video data has seen increased attention recently as it allows training of semantically meaningful embeddings without human annotation, enabling tasks like zero-shot retrieval and action localization. In this work, we present a multi-modal, modality agnostic fusion transformer that learns to exchange information between multiple modalities, such as video, audio, and text, and integrate them into a fused representation in a joined multi-modal embedding space.

119, TITLE: Swin Transformer V2: Scaling Up Capacity and Resolution
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Swin_Transformer_V2_Scaling_Up_Capacity_and_Resolution_CVPR_2022_paper.html
AUTHORS: Ze Liu, Han Hu, Yutong Lin, Zhuliang Yao, Zhenda Xie, Yixuan Wei, Jia Ning, Yue Cao, Zheng Zhang, Li Dong, Furu Wei, Baining Guo
HIGHLIGHT: We present techniques for scaling Swin Transformer [??] up to 3 billion parameters and making it capable of training with images of up to 1,536x1,536 resolution.

120, TITLE: Neural Template: Topology-Aware Reconstruction and Disentangled Generation of 3D Meshes
http://openaccess.thecvf.com//content/CVPR2022/html/Hui_Neural_Template_Topology-Aware_Reconstruction_and_Disentangled_Generation_of_3D_Meshes_CVPR_2022_paper.html
AUTHORS: Ka-Hei Hui, Ruihui Li, Jingyu Hu, Chi-Wing Fu
HIGHLIGHT: This paper introduces a novel framework called DT-Net for 3D mesh reconstruction and generation via Disentangled Topology.

121, TITLE: DEFEAT: Deep Hidden Feature Backdoor Attacks by Imperceptible Perturbation and Latent Representation Constraints
http://openaccess.thecvf.com//content/CVPR2022/html/Zhao_DEFEAT_Deep_Hidden_Feature_Backdoor_Attacks_by_Imperceptible_Perturbation_and_CVPR_2022_paper.html
AUTHORS: Zhendong Zhao, Xiaojun Chen, Yuexin Xuan, Ye Dong, Dakui Wang, Kaitai Liang
HIGHLIGHT: In this paper, We propose a novel and stealthy backdoor attack - DEFEAT.

122, TITLE: Projective Manifold Gradient Layer for Deep Rotation Regression
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Projective_Manifold_Gradient_Layer_for_Deep_Rotation_Regression_CVPR_2022_paper.html
AUTHORS: Jiayi Chen, Yingda Yin, Tolga Birdal, Baoquan Chen, Leonidas J. Guibas, He Wang
HIGHLIGHT: In this paper, we propose a manifold-aware gradient that directly backpropagates into deep network weights.

123, TITLE: CLIMS: Cross Language Image Matching for Weakly Supervised Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Xie_CLIMS_Cross_Language_Image_Matching_for_Weakly_Supervised_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Jinheng Xie, Xianxu Hou, Kai Ye, Linlin Shen
HIGHLIGHT: In this paper, we propose a novel Cross Language Image Matching (CLIMS) framework, based on the recently introduced Contrastive Language-Image Pre-training (CLIP) model, for WSSS.

124, TITLE: Learning To Refactor Action and Co-Occurrence Features for Temporal Action Localization
http://openaccess.thecvf.com//content/CVPR2022/html/Xia_Learning_To_Refactor_Action_and_Co-Occurrence_Features_for_Temporal_Action_CVPR_2022_paper.html
AUTHORS: Kun Xia, Le Wang, Sanping Zhou, Nanning Zheng, Wei Tang
HIGHLIGHT: In this paper, we explore two orthogonal but complementary aspects of a video snippet, i.e., the action features and the co-occurrence features.

125, TITLE: It's Time for Artistic Correspondence in Music and Video

http://openaccess.thecvf.com//content/CVPR2022/html/Suris_Its_Time_for_Artistic_Correspondence_in_Music_and_Video_CVPR_2022_paper.html
AUTHORS: Dídac Surís, Carl Vondrick, Bryan Russell, Justin Salamon
HIGHLIGHT: We present an approach for recommending a music track for a given video, and vice versa, based on both their temporal alignment and their correspondence at an artistic level.

126, TITLE: Mixed Differential Privacy in Computer Vision
http://openaccess.thecvf.com//content/CVPR2022/html/Golatkar_Mixed_Differential_Privacy_in_Computer_Vision_CVPR_2022_paper.html
AUTHORS: Aditya Golatkar, Alessandro Achille, Yu-Xiang Wang, Aaron Roth, Michael Kearns, Stefano Soatto
HIGHLIGHT: We introduce AdaMix, an adaptive differentially private algorithm for training deep neural network classifiers using both private and public image data.

127, TITLE: AdaFace: Quality Adaptive Margin for Face Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Kim_AdaFace_Quality_Adaptive_Margin_for_Face_Recognition_CVPR_2022_paper.html
AUTHORS: Minchul Kim, Anil K. Jain, Xiaoming Liu
HIGHLIGHT: In this work, we introduce another aspect of adaptiveness in the loss function, namely the image quality.

128, TITLE: Learning Soft Estimator of Keypoint Scale and Orientation With Probabilistic Covariant Loss
http://openaccess.thecvf.com//content/CVPR2022/html/Yan_Learning_Soft_Estimator_of_Keypoint_Scale_and_Orientation_With_Probabilistic_CVPR_2022_paper.html
AUTHORS: Pei Yan, Yihua Tan, Shengzhou Xiong, Yuan Tai, Yansheng Li
HIGHLIGHT: However, hard estimators are difficult to handle the local patches containing structures of different objects or multiple edges. In this paper, a Soft Self-Supervised Estimator (S3Esti) is proposed to overcome this problem by learning to predict multiple scales and orientations.

129, TITLE: DN-DETR: Accelerate DETR Training by Introducing Query DeNoising
http://openaccess.thecvf.com//content/CVPR2022/html/Li_DN-DETR_Accelerate_DETR_Training_by_Introducing_Query_DeNoising_CVPR_2022_paper.html
AUTHORS: Feng Li, Hao Zhang, Shilong Liu, Jian Guo, Lionel M. Ni, Lei Zhang
HIGHLIGHT: We present in this paper a novel denoising training method to speedup DETR (DEtection TRansformer) training and offer a deepened understanding of the slow convergence issue of DETR-like methods.

130, TITLE: HCSC: Hierarchical Contrastive Selective Coding
http://openaccess.thecvf.com//content/CVPR2022/html/Guo_HCSC_Hierarchical_Contrastive_Selective_Coding_CVPR_2022_paper.html
AUTHORS: Yuanfan Guo, Minghao Xu, Jiawen Li, Bingbing Ni, Xuanyu Zhu, Zhenbang Sun, Yi Xu
HIGHLIGHT: In addition, the negative pairs used in these methods are not guaranteed to be semantically distinct, which could further hamper the structural correctness of learned image representations. To tackle these limitations, we propose a novel contrastive learning framework called Hierarchical Contrastive Selective Coding (HCSC).

131, TITLE: TransRank: Self-Supervised Video Representation Learning via Ranking-Based Transformation Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Duan_TransRank_Self-Supervised_Video_Representation_Learning_via_Ranking-Based_Transformation_Recognition_CVPR_2022_paper.html
AUTHORS: Haodong Duan, Nanxuan Zhao, Kai Chen, Dahua Lin
HIGHLIGHT: Based on hard-label classification, existing RecogTrans approaches suffer from noisy supervision signals in pre-training. To mitigate this problem, we developed TransRank, a unified framework for recognizing Transformations in a Ranking formulation.

132, TITLE: KeyTr: Keypoint Transporter for 3D Reconstruction of Deformable Objects in Videos
http://openaccess.thecvf.com//content/CVPR2022/html/Novotny_KeyTr_Keypoint_Transporter_for_3D_Reconstruction_of_Deformable_Objects_in_CVPR_2022_paper.html
AUTHORS: David Novotny, Ignacio Rocco, Samarth Sinha, Alexandre Carlier, Gael Kerchenbaum, Roman Shapovalov, Nikita Smetanin, Natalia Neverova, Benjamin Graham, Andrea Vedaldi
HIGHLIGHT: However, the theory of Non-Rigid Structure from Motion prescribes to constrain the deformations for 3D reconstruction. We thus propose a new model that departs significantly from this prior work.

133, TITLE: Invariant Grounding for Video Question Answering
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Invariant_Grounding_for_Video_Question_Answering_CVPR_2022_paper.html
AUTHORS: Yicong Li, Xiang Wang, Junbin Xiao, Wei Ji, Tat-Seng Chua
HIGHLIGHT: In this work, we first take a causal look at VideoQA and argue that invariant grounding is the key to ruling out the spurious correlations. Towards this end, we propose a new learning framework, Invariant Grounding for VideoQA (IGV), to ground the question-critical scene, whose causal relations with answers are invariant across different interventions on the complement.

134, TITLE: Prompt Distribution Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Lu_Prompt_Distribution_Learning_CVPR_2022_paper.html
AUTHORS: Yuning Lu, Jianzhuang Liu, Yonggang Zhang, Yajing Liu, Xinmei Tian

HIGHLIGHT: We present prompt distribution learning for effectively adapting a pre-trained vision-language model to address downstream recognition tasks.

135, TITLE: RAGO: Recurrent Graph Optimizer for Multiple Rotation Averaging
http://openaccess.thecvf.com//content/CVPR2022/html/Li_RAGO_Recurrent_Graph_Optimizer_for_Multiple_Rotation_Averaging_CVPR_2022_paper.html
AUTHORS: Heng Li, Zhaopeng Cui, Shuaicheng Liu, Ping Tan
HIGHLIGHT: This paper proposes a deep recurrent Rotation Averaging Graph Optimizer (RAGO) for Multiple Rotation Averaging (MRA).

136, TITLE: Arch-Graph: Acyclic Architecture Relation Predictor for Task-Transferable Neural Architecture Search
http://openaccess.thecvf.com//content/CVPR2022/html/Huang_Arch-Graph_Acyclic_Architecture_Relation_Predictor_for_Task-Transferable_Neural_Architecture_Search_CVPR_2022_paper.html
AUTHORS: Minbin Huang, Zhijian Huang, Changlin Li, Xin Chen, Hang Xu, Zhenguo Li, Xiaodan Liang
HIGHLIGHT: In this line of research, effectively modeling task correlations is vital yet highly neglected. Therefore, we propose Arch-Graph, a transferable NAS method that predicts task-specific optimal architectures with respect to given task embeddings.

137, TITLE: On Aliased Resizing and Surprising Subtleties in GAN Evaluation
http://openaccess.thecvf.com//content/CVPR2022/html/Parmar_On_Aliased_Resizing_and_Surprising_Subtleties_in_GAN_Evaluation_CVPR_2022_paper.html
AUTHORS: Gaurav Parmar, Richard Zhang, Jun-Yan Zhu
HIGHLIGHT: This paper shows that choices in low-level image processing have been an under-appreciated aspect of generative modeling.

138, TITLE: Lepard: Learning Partial Point Cloud Matching in Rigid and Deformable Scenes
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Lepard_Learning_Partial_Point_Cloud_Matching_in_Rigid_and_Deformable_CVPR_2022_paper.html
AUTHORS: Yang Li, Tatsuya Harada
HIGHLIGHT: We present Lepard, a Learning based approach for partial point cloud matching in rigid and deformable scenes.

139, TITLE: Virtual Elastic Objects
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Virtual_Elastic_Objects_CVPR_2022_paper.html
AUTHORS: Hsiao-yu Chen, Edith Tretschk, Tuur Stuyck, Petr Kadlecek, Ladislav Kavan, Etienne Vouga, Christoph Lassner
HIGHLIGHT: We present Virtual Elastic Objects (VEOs): virtual objects that not only look like their real-world counterparts but also behave like them, even when subject to novel interactions.

140, TITLE: DiSparse: Disentangled Sparsification for Multitask Model Compression
http://openaccess.thecvf.com//content/CVPR2022/html/Sun_DiSparse_Disentangled_Sparsification_for_Multitask_Model_Compression_CVPR_2022_paper.html
AUTHORS: Xinglong Sun, Ali Hassani, Zhangyang Wang, Gao Huang, Humphrey Shi
HIGHLIGHT: In this paper, we propose DiSparse, a simple, effective, and first-of-its-kind multitask pruning and sparse training scheme.

141, TITLE: Pushing the Limits of Simple Pipelines for Few-Shot Learning: External Data and Fine-Tuning Make a Difference
http://openaccess.thecvf.com//content/CVPR2022/html/Hu_Pushing_the_Limits_of_Simple_Pipelines_for_Few-Shot_Learning_External_CVPR_2022_paper.html
AUTHORS: Shell Xu Hu, Da Li, Jan Stühmer, Minyoung Kim, Timothy M. Hospedales
HIGHLIGHT: We seek to push the limits of a simple-but-effective pipeline for real-world few-shot image classification in practice. To this end, we explore few-shot learning from the perspective of neural architecture, as well as a three stage pipeline of pre-training on external data, meta-training with labelled few-shot tasks, and task-specific fine-tuning on unseen tasks.

142, TITLE: Opening Up Open World Tracking
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Opening_Up_Open_World_Tracking_CVPR_2022_paper.html
AUTHORS: Yang Liu, Idil Esen Zulfikar, Jonathon Luiten, Achal Dave, Deva Ramanan, Bastian Leibe, Aljoša Ošep, Laura Leal-Taixé
HIGHLIGHT: This paper addresses this evaluation deficit and lays out the landscape and evaluation methodology for detecting and tracking both known and unknown objects in the open-world setting.

143, TITLE: Towards Efficient and Scalable Sharpness-Aware Minimization
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Towards_Efficient_and_Scalable_Sharpness-Aware_Minimization_CVPR_2022_paper.html
AUTHORS: Yong Liu, Siqi Mai, Xiangning Chen, Cho-Jui Hsieh, Yang You
HIGHLIGHT: In this paper, we propose a novel algorithm LookSAM - that only periodically calculates the inner gradient ascent, to significantly reduce the additional training cost of SAM.

144, TITLE: VISTA: Boosting 3D Object Detection via Dual Cross-VIew SpaTial Attention

http://openaccess.thecvf.com//content/CVPR2022/html/Deng_VISTA_Boosting_3D_Object_Detection_via_Dual_Cross-VIew_SpaTial_Attention_CVPR_2022_paper.html
AUTHORS: Shengheng Deng, Zhihao Liang, Lin Sun, Kui Jia
HIGHLIGHT: In this paper, we propose to adaptively fuse multi-view features in a global spatial context via Dual Cross-VIew SpaTial Attention (VISTA).

145, TITLE: Rethinking Deep Face Restoration
http://openaccess.thecvf.com//content/CVPR2022/html/Zhao_Rethinking_Deep_Face_Restoration_CVPR_2022_paper.html
AUTHORS: Yang Zhao, Yu-Chuan Su, Chun-Te Chu, Yandong Li, Marius Renn, Yukun Zhu, Changyou Chen, Xuhui Jia
HIGHLIGHT: Because the human visual system is very sensitive to faces, even minor facial changes may alter the identity and significantly degrade the perceptual quality. In this work, we argue the problems of existing models can be traced down to the two sub-tasks of the face restoration problem, i.e. face generation and face reconstruction, and the fragile balance between them.

146, TITLE: OSSO: Obtaining Skeletal Shape From Outside
http://openaccess.thecvf.com//content/CVPR2022/html/Keller_OSSO_Obtaining_Skeletal_Shape_From_Outside_CVPR_2022_paper.html
AUTHORS: Marilyn Keller, Silvia Zuffi, Michael J. Black, Sergi Pujades
HIGHLIGHT: We address the problem of inferring the anatomic skeleton of a person, in an arbitrary pose, from the 3D surface of the body; i.e. we predict the inside (bones) from the outside (skin).

147, TITLE: Temporal Alignment Networks for Long-Term Video
http://openaccess.thecvf.com//content/CVPR2022/html/Han_Temporal_Alignment_Networks_for_Long-Term_Video_CVPR_2022_paper.html
AUTHORS: Tengda Han, Weidi Xie, Andrew Zisserman
HIGHLIGHT: The objective of this paper is a temporal alignment network that ingests long term video sequences, and associated text sentences, in order to: (1) determine if a sentence is alignable with the video; and (2) if it is alignable, then determine its alignment.

148, TITLE: Few-Shot Head Swapping in the Wild
http://openaccess.thecvf.com//content/CVPR2022/html/Shu_Few-Shot_Head_Swapping_in_the_Wild_CVPR_2022_paper.html
AUTHORS: Changyong Shu, Hemao Wu, Hang Zhou, Jiaming Liu, Zhibin Hong, Changxing Ding, Junyu Han, Jingtuo Liu, Errui Ding, Jingdong Wang
HIGHLIGHT: In this paper, we present the Head Swapper (HeSer), which achieves few-shot head swapping in the wild through two dedicated designed modules.

149, TITLE: A Study on the Distribution of Social Biases in Self-Supervised Learning Visual Models
http://openaccess.thecvf.com//content/CVPR2022/html/Sirotkin_A_Study_on_the_Distribution_of_Social_Biases_in_Self-Supervised_CVPR_2022_paper.html
AUTHORS: Kirill Sirotkin, Pablo Carballeira, Marcos Escudero-Viñolo
HIGHLIGHT: In this paper, we study the biases of a varied set of SSL visual models, trained using ImageNet data, using a method and dataset designed by psychological experts to measure social biases.

150, TITLE: LAR-SR: A Local Autoregressive Model for Image Super-Resolution
http://openaccess.thecvf.com//content/CVPR2022/html/Guo_LAR-SR_A_Local_Autoregressive_Model_for_Image_Super-Resolution_CVPR_2022_paper.html
AUTHORS: Baisong Guo, Xiaoyun Zhang, Haoning Wu, Yu Wang, Ya Zhang, Yan-Feng Wang
HIGHLIGHT: Based on the fact that given the structural information, the textural details in the natural images are locally related without long term dependency, in this paper we propose a novel autoregressive model-based SR approach, namely LAR-SR, which can efficiently generate realistic SR images using a novel local autoregressive (LAR) module.

151, TITLE: Bayesian Invariant Risk Minimization
http://openaccess.thecvf.com//content/CVPR2022/html/Lin_Bayesian_Invariant_Risk_Minimization_CVPR_2022_paper.html
AUTHORS: Yong Lin, Hanze Dong, Hao Wang, Tong Zhang
HIGHLIGHT: Our empirical evidence also provides supports: IRM methods that work well in typical settings significantly deteriorate even if we slightly enlarge the model size or lessen the training data. To alleviate this issue, we propose Bayesian Invariant Risk Minimization (BIRM) by introducing Bayesian inference into the IRM.

152, TITLE: Democracy Does Matter: Comprehensive Feature Mining for Co-Salient Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Yu_Democracy_Does_Matter_Comprehensive_Feature_Mining_for_Co-Salient_Object_Detection_CVPR_2022_paper.html
AUTHORS: Siyue Yu, Jimin Xiao, Bingfeng Zhang, Eng Gee Lim
HIGHLIGHT: In this paper, we aim to mine comprehensive co-salient features with democracy and reduce background interference without introducing any extra information.

153, TITLE: Alleviating Semantics Distortion in Unsupervised Low-Level Image-to-Image Translation via Structure Consistency Constraint
http://openaccess.thecvf.com//content/CVPR2022/html/Guo_Alleviating_Semantics_Distortion_in_Unsupervised_Low-Level_Image-to-Image_Translation_via_Structure_CVPR_2022_paper.html
AUTHORS: Jiaxian Guo, Jiachen Li, Huan Fu, Mingming Gong, Kun Zhang, Dacheng Tao

HIGHLIGHT: In this paper, we focus on the low-level I2I translation, where the structure of images is highly related to their semantics.

154, TITLE: Doodle It Yourself: Class Incremental Learning by Drawing a Few Sketches
http://openaccess.thecvf.com//content/CVPR2022/html/Bhunia_Doodle_It_Yourself_Class_Incremental_Learning_by_Drawing_a_Few_CVPR_2022_paper.html
AUTHORS: Ayan Kumar Bhunia, Viswanatha Reddy Gajjala, Subhadeep Koley, Rohit Kundu, Aneeshan Sain, Tao Xiang, Yi-Zhe Song
HIGHLIGHT: For that, we present a framework that infuses (i) gradient consensus for domain invariant learning, (ii) knowledge distillation for preserving old class information, and (iii) graph attention networks for message passing between old and novel classes.

155, TITLE: Self-Supervised Predictive Learning: A Negative-Free Method for Sound Source Localization in Visual Scenes
http://openaccess.thecvf.com//content/CVPR2022/html/Song_Self-Supervised_Predictive_Learning_A_Negative-Free_Method_for_Sound_Source_Localization_CVPR_2022_paper.html
AUTHORS: Zengjie Song, Yuxi Wang, Junsong Fan, Tieniu Tan, Zhaoxiang Zhang
HIGHLIGHT: In this paper, instead of following previous literature, we propose Self-Supervised Predictive Learning (SSPL), a negative-free method for sound localization via explicit positive mining.

156, TITLE: ICON: Implicit Clothed Humans Obtained From Normals
http://openaccess.thecvf.com//content/CVPR2022/html/Xiu_ICON_Implicit_Clothed_Humans_Obtained_From_Normals_CVPR_2022_paper.html
AUTHORS: Yuliang Xiu, Jinlong Yang, Dimitrios Tzionas, Michael J. Black
HIGHLIGHT: In contrast, our goal is to learn the avatar from only 2D images of people in unconstrained poses.

157, TITLE: Comparing Correspondences: Video Prediction With Correspondence-Wise Losses
http://openaccess.thecvf.com//content/CVPR2022/html/Geng_Comparing_Correspondences_Video_Prediction_With_Correspondence-Wise_Losses_CVPR_2022_paper.html
AUTHORS: Daniel Geng, Max Hamilton, Andrew Owens
HIGHLIGHT: Image prediction methods often struggle on tasks that require changing the positions of objects, such as video prediction, producing blurry images that average over the many positions that objects might occupy. In this paper, we propose a simple change to existing image similarity metrics that makes them more robust to positional errors: we match the images using optical flow, then measure the visual similarity of corresponding pixels.

158, TITLE: Uni-Perceiver: Pre-Training Unified Architecture for Generic Perception for Zero-Shot and Few-Shot Tasks
http://openaccess.thecvf.com//content/CVPR2022/html/Zhu_Uni-Perceiver_Pre-Training_Unified_Architecture_for_Generic_Perception_for_Zero-Shot_and_CVPR_2022_paper.html
AUTHORS: Xizhou Zhu, Jinguo Zhu, Hao Li, Xiaoshi Wu, Hongsheng Li, Xiaohua Wang, Jifeng Dai
HIGHLIGHT: In this paper, we present a generic perception architecture named Uni-Perceiver, which processes a variety of modalities and tasks with unified modeling and shared parameters.

159, TITLE: The Auto Arborist Dataset: A Large-Scale Benchmark for Multiview Urban Forest Monitoring Under Domain Shift
http://openaccess.thecvf.com//content/CVPR2022/html/Beery_The_Auto_Arborist_Dataset_A_Large-Scale_Benchmark_for_Multiview_Urban_CVPR_2022_paper.html
AUTHORS: Sara Beery, Guanhang Wu, Trevor Edwards, Filip Pavetic, Bo Majewski, Shreyasee Mukherjee, Stanley Chan, John Morgan, Vivek Rathod, Jonathan Huang
HIGHLIGHT: To build computer vision systems that truly solve real-world problems at global scale, we need benchmarks that fully capture real-world complexity, including geographic domain shift, long-tailed distributions, and data noise. We propose urban forest monitoring as an ideal testbed for studying and improving upon these computer vision challenges, while simultaneously working towards filling a crucial environmental and societal need.

160, TITLE: On the Instability of Relative Pose Estimation and RANSAC's Role
http://openaccess.thecvf.com//content/CVPR2022/html/Fan_On_the_Instability_of_Relative_Pose_Estimation_and_RANSACs_Role_CVPR_2022_paper.html
AUTHORS: Hongyi Fan, Joe Kileel, Benjamin Kimia
HIGHLIGHT: These cases arise due to numerical instability of the 5- and 7-point minimal problems. This paper characterizes these instabilities, both in terms of minimal world scene configurations that lead to infinite condition number in epipolar estimation, and also in terms of the related minimal image feature pair correspondence configurations.

161, TITLE: Shape From Polarization for Complex Scenes in the Wild
http://openaccess.thecvf.com//content/CVPR2022/html/Lei_Shape_From_Polarization_for_Complex_Scenes_in_the_Wild_CVPR_2022_paper.html
AUTHORS: Chenyang Lei, Chenyang Qi, Jiaxin Xie, Na Fan, Vladlen Koltun, Qifeng Chen
HIGHLIGHT: We present a new data-driven approach with physics-based priors to scene-level normal estimation from a single polarization image.

162, TITLE: Real-Time, Accurate, and Consistent Video Semantic Segmentation via Unsupervised Adaptation and Cross-Unit Deployment on Mobile Device

http://openaccess.thecvf.com//content/CVPR2022/html/Park_Real-Time_Accurate_and_Consistent_Video_Semantic_Segmentation_via_Unsupervised_Adaptation_CVPR_2022_paper.html
AUTHORS:		Hyojin Park, Alan Yessenbayev, Tushar Singhal, Navin Kumar Adhikari, Yizhe Zhang, Shubhankar Mangesh Borse, Hong Cai, Nilesh Prasad Pandey, Fei Yin, Frank Mayer, Balaji Calidas, Fatih Porikli
HIGHLIGHT:		This demonstration showcases our innovations on efficient, accurate, and temporally consistent video semantic segmentation on mobile device.

163, TITLE:		SNUG: Self-Supervised Neural Dynamic Garments
http://openaccess.thecvf.com//content/CVPR2022/html/Santesteban_SNUG_Self-Supervised_Neural_Dynamic_Garments_CVPR_2022_paper.html
AUTHORS:		Igor Santesteban, Miguel A. Otaduy, Dan Casas
HIGHLIGHT:		We present a self-supervised method to learn dynamic 3D deformations of garments worn by parametric human bodies.

164, TITLE:		Towards Fewer Annotations: Active Learning via Region Impurity and Prediction Uncertainty for Domain Adaptive Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Xie_Towards_Fewer_Annotations_Active_Learning_via_Region_Impurity_and_Prediction_CVPR_2022_paper.html
AUTHORS:		Binhui Xie, Longhui Yuan, Shuang Li, Chi Harold Liu, Xinjing Cheng
HIGHLIGHT:		In this paper, we propose a simple region-based active learning approach for semantic segmentation under a domain shift, aiming to automatically query a small partition of image regions to be labeled while maximizing segmentation performance.

165, TITLE:		Glass Segmentation Using Intensity and Spectral Polarization Cues
http://openaccess.thecvf.com//content/CVPR2022/html/Mei_Glass_Segmentation_Using_Intensity_and_Spectral_Polarization_Cues_CVPR_2022_paper.html
AUTHORS:		Haiyang Mei, Bo Dong, Wen Dong, Jiaxi Yang, Seung-Hwan Baek, Felix Heide, Pieter Peers, Xiaopeng Wei, Xin Yang
HIGHLIGHT:		In this work, we exploit that the light-matter interactions on glass materials provide unique intensity-polarization cues for each observed wavelength of light.

166, TITLE:		CrossPoint: Self-Supervised Cross-Modal Contrastive Learning for 3D Point Cloud Understanding
http://openaccess.thecvf.com//content/CVPR2022/html/Afham_CrossPoint_Self-Supervised_Cross-Modal_Contrastive_Learning_for_3D_Point_Cloud_Understanding_CVPR_2022_paper.html
AUTHORS:		Mohamed Afham, Isuru Dissanayake, Dinithi Dissanayake, Amaya Dharmasiri, Kanchana Thilakarathna, Ranga Rodrigo
HIGHLIGHT:		We observe in the real world that humans are capable of mapping the visual concepts learnt from 2D images to understand the 3D world. Encouraged by this insight, we propose CrossPoint, a simple cross-modal contrastive learning approach to learn transferable 3D point cloud representations.

167, TITLE:		Few Shot Generative Model Adaption via Relaxed Spatial Structural Alignment
http://openaccess.thecvf.com//content/CVPR2022/html/Xiao_Few_Shot_Generative_Model_Adaption_via_Relaxed_Spatial_Structural_Alignment_CVPR_2022_paper.html
AUTHORS:		Jiayu Xiao, Liang Li, Chaofei Wang, Zheng-Jun Zha, Qingming Huang
HIGHLIGHT:		However, existing methods are prone to model overfitting and collapse in extremely few shot setting (less than 10). To solve this problem, we propose a relaxed spatial structural alignment (RSSA) method to calibrate the target generative models during the adaption.

168, TITLE:		Target-Relevant Knowledge Preservation for Multi-Source Domain Adaptive Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Wu_Target-Relevant_Knowledge_Preservation_for_Multi-Source_Domain_Adaptive_Object_Detection_CVPR_2022_paper.html
AUTHORS:		Jiaxi Wu, Jiaxin Chen, Mengzhe He, Yiru Wang, Bo Li, Bingqi Ma, Weihao Gan, Wei Wu, Yali Wang, Di Huang
HIGHLIGHT:		Albeit great effort made in single source domain adaptation, a more generalized task with multiple source domains remains not being well explored, due to knowledge degradation during their combination. To address this issue, we propose a novel approach, namely target-relevant knowledge preservation (TRKP), to unsupervised multi-source DAOD.

169, TITLE:		Pyramid Grafting Network for One-Stage High Resolution Saliency Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Xie_Pyramid_Grafting_Network_for_One-Stage_High_Resolution_Saliency_Detection_CVPR_2022_paper.html
AUTHORS:		Chenxi Xie, Changqun Xia, Mingcan Ma, Zhirui Zhao, Xiaowu Chen, Jia Li
HIGHLIGHT:		However, most of existing SOD models designed for low-resolution input perform poorly on high-resolution images due to the contradiction between the sampling depth and the receptive field size. Aiming at resolving this contradiction, we propose a novel one-stage framework called Pyramid Grafting Network (PGNet), using transformer and CNN backbone to extract features from different resolution images independently and then graft the features from transformer branch to CNN branch.

170, TITLE:		A Style-Aware Discriminator for Controllable Image Translation
http://openaccess.thecvf.com//content/CVPR2022/html/Kim_A_Style-Aware_Discriminator_for_Controllable_Image_Translation_CVPR_2022_paper.html
AUTHORS:		Kunhee Kim, Sanghun Park, Eunyeong Jeon, Taehun Kim, Daijin Kim

HIGHLIGHT: This limitation largely arises because labels do not consider the semantic distance. To mitigate such problems, we propose a style-aware discriminator that acts as a critic as well as a style encoder to provide conditions.

171, TITLE: Non-Iterative Recovery From Nonlinear Observations Using Generative Models
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Non-Iterative_Recovery_From_Nonlinear_Observations_Using_Generative_Models_CVPR_2022_paper.html
AUTHORS: Jiulong Liu, Zhaoqiang Liu
HIGHLIGHT: In this paper, we aim to estimate the direction of an underlying signal from its nonlinear observations following the semi-parametric single index model (SIM).

172, TITLE: Incremental Cross-View Mutual Distillation for Self-Supervised Medical CT Synthesis
http://openaccess.thecvf.com//content/CVPR2022/html/Fang_Incremental_Cross-View_Mutual_Distillation_for_Self-Supervised_Medical_CT_Synthesis_CVPR_2022_paper.html
AUTHORS: Chaowei Fang, Liang Wang, Dingwen Zhang, Jun Xu, Yixuan Yuan, Junwei Han
HIGHLIGHT: To this end, this paper builds a novel medical slice synthesis to increase the inter-slice resolution. Considering that the ground-truth intermediate medical slices are always absent in clinical practice, we introduce the incremental cross-view mutual distillation strategy to accomplish this task in the self-supervised learning manner.

173, TITLE: Enhancing Adversarial Training With Second-Order Statistics of Weights
http://openaccess.thecvf.com//content/CVPR2022/html/Jin_Enhancing_Adversarial_Training_With_Second-Order_Statistics_of_Weights_CVPR_2022_paper.html
AUTHORS: Gaojie Jin, Xinping Yi, Wei Huang, Sven Schewe, Xiaowei Huang
HIGHLIGHT: In this paper, we show that treating model weights as random variables allows for enhancing adversarial training through Second-Order Statistics Optimization (S^2O) with respect to the weights.

174, TITLE: Partially Does It: Towards Scene-Level FG-SBIR With Partial Input
http://openaccess.thecvf.com//content/CVPR2022/html/Chowdhury_Partially_Does_It_Towards_Scene-Level_FG-SBIR_With_Partial_Input_CVPR_2022_paper.html
AUTHORS: Pinaki Nath Chowdhury, Ayan Kumar Bhunia, Viswanatha Reddy Gajjala, Aneeshan Sain, Tao Xiang, Yi-Zhe Song
HIGHLIGHT: A quick pilot study reveals: (i) a scene sketch does not necessarily contain all objects in the corresponding photo, due to the subjective holistic interpretation of scenes, (ii) there exists significant empty (white) regions as a result of object-level abstraction, and as a result, (iii) existing scene-level fine-grained sketch-based image retrieval methods collapse as scene sketches become more partial. To solve this "partial" problem, we advocate for a simple set-based approach using optimal transport (OT) to model cross-modal region associativity in a partially-aware fashion.

175, TITLE: Dual Temperature Helps Contrastive Learning Without Many Negative Samples: Towards Understanding and Simplifying MoCo
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Dual_Temperature_Helps_Contrastive_Learning_Without_Many_Negative_Samples_Towards_CVPR_2022_paper.html
AUTHORS: Chaoning Zhang, Kang Zhang, Trung X. Pham, Axi Niu, Zhinan Qiao, Chang D. Yoo, In So Kweon
HIGHLIGHT: Our findings motivate us to simplify MoCo v2 via the removal of its dictionary as well as momentum.

176, TITLE: Moving Window Regression: A Novel Approach to Ordinal Regression
http://openaccess.thecvf.com//content/CVPR2022/html/Shin_Moving_Window_Regression_A_Novel_Approach_to_Ordinal_Regression_CVPR_2022_paper.html
AUTHORS: Nyeong-Ho Shin, Seon-Ho Lee, Chang-Su Kim
HIGHLIGHT: A novel ordinal regression algorithm, called moving window regression (MWR), is proposed in this paper.

177, TITLE: UniCoRN: A Unified Conditional Image Repainting Network
http://openaccess.thecvf.com//content/CVPR2022/html/Sun_UniCoRN_A_Unified_Conditional_Image_Repainting_Network_CVPR_2022_paper.html
AUTHORS: Jimeng Sun, Shuchen Weng, Zheng Chang, Si Li, Boxin Shi
HIGHLIGHT: Existing methods based on two-phase architecture design assume dependency between phases and cause color-image incongruity. To solve these problems, we propose a novel Unified Conditional image Repainting Network (UniCoRN).

178, TITLE: Forecasting Characteristic 3D Poses of Human Actions
http://openaccess.thecvf.com//content/CVPR2022/html/Diller_Forecasting_Characteristic_3D_Poses_of_Human_Actions_CVPR_2022_paper.html
AUTHORS: Christian Diller, Thomas Funkhouser, Angela Dai
HIGHLIGHT: To predict characteristic poses, we propose a probabilistic approach that models the possible multi-modality in the distribution of likely characteristic poses.

179, TITLE: ACPL: Anti-Curriculum Pseudo-Labelling for Semi-Supervised Medical Image Classification
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_ACPL_Anti-Curriculum_Pseudo-Labelling_for_Semi-Supervised_Medical_Image_Classification_CVPR_2022_paper.html
AUTHORS: Fengbei Liu, Yu Tian, Yuanhong Chen, Yuyuan Liu, Vasileios Belagiannis, Gustavo Carneiro
HIGHLIGHT: In this paper, unlike traditional methods that select confident pseudo label by threshold, we propose a new SSL algorithm, called anti-curriculum pseudo-labelling (ACPL), which introduces novel techniques to select informative unlabelled

samples, improving training balance and allowing the model to work for both multi-label and multi-class problems, and to estimate pseudo labels by an accurate ensemble of classifiers (improving pseudo label accuracy).

180, TITLE: Learning to Deblur Using Light Field Generated and Real Defocus Images
http://openaccess.thecvf.com//content/CVPR2022/html/Ruan_Learning_to_Deblur_Using_Light_Field_Generated_and_Real_Defocus_CVPR_2022_paper.html
AUTHORS: Lingyan Ruan, Bin Chen, Jizhou Li, Miuling Lam
HIGHLIGHT: In this paper, we propose a novel deep defocus deblurring network that leverages the strength and overcomes the shortcoming of light fields.

181, TITLE: Self-Supervised Predictive Convolutional Attentive Block for Anomaly Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Ristea_Self-Supervised_Predictive_Convolutional_Attentive_Block_for_Anomaly_Detection_CVPR_2022_paper.html
AUTHORS: Nicolae-Cătălin Ristea, Neelu Madan, Radu Tudor Ionescu, Kamal Nasrollahi, Fahad Shahbaz Khan, Thomas B. Moeslund, Mubarak Shah
HIGHLIGHT: Different from related methods, we propose to integrate the reconstruction-based functionality into a novel self-supervised predictive architectural building block.

182, TITLE: Safe Self-Refinement for Transformer-Based Domain Adaptation
http://openaccess.thecvf.com//content/CVPR2022/html/Sun_Safe_Self-Refinement_for_Transformer-Based_Domain_Adaptation_CVPR_2022_paper.html
AUTHORS: Tao Sun, Cheng Lu, Tianshuo Zhang, Haibin Ling
HIGHLIGHT: In this paper we propose a novel solution named SSRT (Safe Self-Refinement for Transformer-based domain adaptation), which brings improvement from two aspects.

183, TITLE: Density-Preserving Deep Point Cloud Compression
http://openaccess.thecvf.com//content/CVPR2022/html/He_Density-Preserving_Deep_Point_Cloud_Compression_CVPR_2022_paper.html
AUTHORS: Yun He, Xinlin Ren, Danhang Tang, Yinda Zhang, Xiangyang Xue, Yanwei Fu
HIGHLIGHT: Local density of point clouds is crucial for representing local details, but has been overlooked by existing point cloud compression methods. To address this, we propose a novel deep point cloud compression method that preserves local density information.

184, TITLE: StyleMesh: Style Transfer for Indoor 3D Scene Reconstructions
http://openaccess.thecvf.com//content/CVPR2022/html/Hollein_StyleMesh_Style_Transfer_for_Indoor_3D_Scene_Reconstructions_CVPR_2022_paper.html
AUTHORS: Lukas Höllein, Justin Johnson, Matthias Nießner
HIGHLIGHT: We apply style transfer on mesh reconstructions of indoor scenes.

185, TITLE: Which Model To Transfer? Finding the Needle in the Growing Haystack
http://openaccess.thecvf.com//content/CVPR2022/html/Renggli_Which_Model_To_Transfer_Finding_the_Needle_in_the_Growing_CVPR_2022_paper.html
AUTHORS: Cedric Renggli, André Susano Pinto, Luka Rimanic, Joan Puigcerver, Carlos Riquelme, Ce Zhang, Mario Lučić
HIGHLIGHT: As these repositories keep growing exponentially, efficiently selecting a good model for the task at hand becomes paramount. We provide a formalization of this problem through a familiar notion of regret and introduce the predominant strategies, namely task-agnostic (e.g. ranking models by their ImageNet performance) and task-aware search strategies (such as linear or kNN evaluation).

186, TITLE: Fast and Unsupervised Action Boundary Detection for Action Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Du_Fast_and_Unsupervised_Action_Boundary_Detection_for_Action_Segmentation_CVPR_2022_paper.html
AUTHORS: Zexing Du, Xue Wang, Guoqing Zhou, Qing Wang
HIGHLIGHT: To deal with the great number of untrimmed videos produced every day, we propose an efficient unsupervised action segmentation method by detecting boundaries, named action boundary detection (ABD).

187, TITLE: Class-Incremental Learning With Strong Pre-Trained Models
http://openaccess.thecvf.com//content/CVPR2022/html/Wu_Class-Incremental_Learning_With_Strong_Pre-Trained_Models_CVPR_2022_paper.html
AUTHORS: Tz-Ying Wu, Gurumurthy Swaminathan, Zhizhong Li, Avinash Ravichandran, Nuno Vasconcelos, Rahul Bhotika, Stefano Soatto
HIGHLIGHT: We propose a 2-stage training scheme, i) feature augmentation - cloning part of the backbone and fine-tuning it on the novel data, and ii) fusion - combining the base and novel classifiers into a unified classifier.

188, TITLE: Robust Optimization As Data Augmentation for Large-Scale Graphs
http://openaccess.thecvf.com//content/CVPR2022/html/Kong_Robust_Optimization_As_Data_Augmentation_for_Large-Scale_Graphs_CVPR_2022_paper.html
AUTHORS: Kezhi Kong, Guohao Li, Mucong Ding, Zuxuan Wu, Chen Zhu, Bernard Ghanem, Gavin Taylor, Tom Goldstein

HIGHLIGHT: We propose FLAG (Free Large-scale Adversarial Augmentation on Graphs), which iteratively augments node features with gradient-based adversarial perturbations during training.

189, TITLE: Robust Structured Declarative Classifiers for 3D Point Clouds: Defending Adversarial Attacks With Implicit Gradients
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Robust_Structured_Declarative_Classifiers_for_3D_Point_Clouds_Defending_Adversarial_CVPR_2022_paper.html
AUTHORS: Kaidong Li, Ziming Zhang, Cuncong Zhong, Guanghui Wang
HIGHLIGHT: In contrast to the literature, we propose a family of robust structured declarative classifiers for point cloud classification, where the internal constrained optimization mechanism can effectively defend adversarial attacks through implicit gradients.

190, TITLE: PhotoScene: Photorealistic Material and Lighting Transfer for Indoor Scenes
http://openaccess.thecvf.com//content/CVPR2022/html/Yeh_PhotoScene_Photorealistic_Material_and_Lighting_Transfer_for_Indoor_Scenes_CVPR_2022_paper.html
AUTHORS: Yu-Ying Yeh, Zhengqin Li, Yannick Hold-Geoffroy, Rui Zhu, Zexiang Xu, MiloÅ¡ HaÅ¡an, Kalyan Sunkavalli, Manmohan Chandraker
HIGHLIGHT: Most indoor 3D scene reconstruction methods focus on recovering 3D geometry and scene layout. In this work, we go beyond this to propose PhotoScene, a framework that takes input image(s) of a scene along with approximately aligned CAD geometry (either reconstructed automatically or manually specified) and builds a photorealistic digital twin with high-quality materials and similar lighting.

191, TITLE: Improving the Transferability of Targeted Adversarial Examples Through Object-Based Diverse Input
http://openaccess.thecvf.com//content/CVPR2022/html/Byun_Improving_the_Transferability_of_Targeted_Adversarial_Examples_Through_Object-Based_Diverse_CVPR_2022_paper.html
AUTHORS: Junyoung Byun, Seungju Cho, Myung-Joon Kwon, Hee-Seon Kim, Changick Kim
HIGHLIGHT: However, prior works utilize simple image transformations such as resizing, which limits input diversity. To tackle this limitation, we propose the object-based diverse input (ODI) method that draws an adversarial image on a 3D object and induces the rendered image to be classified as the target class.

192, TITLE: IRON: Inverse Rendering by Optimizing Neural SDFs and Materials From Photometric Images
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_IRON_Inverse_Rendering_by_Optimizing_Neural_SDFs_and_Materials_From_CVPR_2022_paper.html
AUTHORS: Kai Zhang, Fujun Luan, Zhengqi Li, Noah Snavely
HIGHLIGHT: We propose a neural inverse rendering pipeline called IRON that operates on photometric images and outputs high-quality 3D content in the format of triangle meshes and material textures readily deployable in existing graphics pipelines.

193, TITLE: ObjectFolder 2.0: A Multisensory Object Dataset for Sim2Real Transfer
http://openaccess.thecvf.com//content/CVPR2022/html/Gao_ObjectFolder_2.0_A_Multisensory_Object_Dataset_for_Sim2Real_Transfer_CVPR_2022_paper.html
AUTHORS: Ruohan Gao, Zilin Si, Yen-Yu Chang, Samuel Clarke, Jeannette Bohg, Li Fei-Fei, Wenzhen Yuan, Jiajun Wu
HIGHLIGHT: We present ObjectFolder 2.0, a large-scale, multisensory dataset of common household objects in the form of implicit neural representations that significantly enhances ObjectFolder 1.0 in three aspects.

194, TITLE: Versatile Multi-Modal Pre-Training for Human-Centric Perception
http://openaccess.thecvf.com//content/CVPR2022/html/Hong_Versatile_Multi-Modal_Pre-Training_for_Human-Centric_Perception_CVPR_2022_paper.html
AUTHORS: Fangzhou Hong, Liang Pan, Zhongang Cai, Ziwei Liu
HIGHLIGHT: Therefore, it is desirable to have a versatile pre-train model that serves as a foundation for data-efficient downstream tasks transfer. To this end, we propose the Human-Centric Multi-Modal Contrastive Learning framework HCMoCo that leverages the multi-modal nature of human data (e.g. RGB, depth, 2D keypoints) for effective representation learning.

195, TITLE: 360MonoDepth: High-Resolution 360deg Monocular Depth Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Rey-Area_360MonoDepth_High-Resolution_360deg_Monocular_Depth_Estimation_CVPR_2022_paper.html
AUTHORS: Manuel Rey-Area, Mingze Yuan, Christian Richardt
HIGHLIGHT: In this work, we propose a flexible framework for monocular depth estimation from high-resolution 360deg images using tangent images.

196, TITLE: Splicing ViT Features for Semantic Appearance Transfer
http://openaccess.thecvf.com//content/CVPR2022/html/Tumanyan_Splicing_ViT_Features_for_Semantic_Appearance_Transfer_CVPR_2022_paper.html
AUTHORS: Narek Tumanyan, Omer Bar-Tal, Shai Bagon, Tali Dekel
HIGHLIGHT: We present a method for semantically transferring the visual appearance of one natural image to another.

197, TITLE: Contrastive Regression for Domain Adaptation on Gaze Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Contrastive_Regression_for_Domain_Adaptation_on_Gaze_Estimation_CVPR_2022_paper.html
AUTHORS: Yaoming Wang, Yangzhou Jiang, Jin Li, Bingbing Ni, Wenrui Dai, Chenglin Li, Hongkai Xiong, Teng Li

HIGHLIGHT: In this paper, we propose a novel gaze adaptation approach, namely Contrastive Regression Gaze Adaptation (CRGA), for generalizing gaze estimation on the target domain in an unsupervised manner.

198, TITLE: MUSE-VAE: Multi-Scale VAE for Environment-Aware Long Term Trajectory Prediction
http://openaccess.thecvf.com//content/CVPR2022/html/Lee_MUSE-VAE_Multi-Scale_VAE_for_Environment-Aware_Long_Term_Trajectory_Prediction_CVPR_2022_paper.html
AUTHORS: Mihee Lee, Samuel S. Sohn, Seonghyeon Moon, Sejong Yoon, Mubbasir Kapadia, Vladimir Pavlovic
HIGHLIGHT: In this work, we propose MUSE-VAE, a new probabilistic modeling framework based on a cascade of Conditional VAEs, which tackles the long-term, uncertain trajectory prediction task using a coarse-to-fine multi-factor forecasting architecture.

199, TITLE: Multi-View Consistent Generative Adversarial Networks for 3D-Aware Image Synthesis
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Multi-View_Consistent_Generative_Adversarial_Networks_for_3D-Aware_Image_Synthesis_CVPR_2022_paper.html
AUTHORS: Xuanmeng Zhang, Zhedong Zheng, Daiheng Gao, Bang Zhang, Pan Pan, Yi Yang
HIGHLIGHT: However, one key challenge remains: existing approaches lack geometry constraints, hence usually fail to generate multi-view consistent images. To address this challenge, we propose Multi-View Consistent Generative Adversarial Networks (MVCGAN) for high-quality 3D-aware image synthesis with geometry constraints.

200, TITLE: Putting People in Their Place: Monocular Regression of 3D People in Depth
http://openaccess.thecvf.com//content/CVPR2022/html/Sun_Putting_People_in_Their_Place_Monocular_Regression_of_3D_People_CVPR_2022_paper.html
AUTHORS: Yu Sun, Wu Liu, Qian Bao, Yili Fu, Tao Mei, Michael J. Black
HIGHLIGHT: Given an image with multiple people, our goal is to directly regress the pose and shape of all the people as well as their relative depth.

201, TITLE: POCO: Point Convolution for Surface Reconstruction
http://openaccess.thecvf.com//content/CVPR2022/html/Boulch_POCO_Point_Convolution_for_Surface_Reconstruction_CVPR_2022_paper.html
AUTHORS: Alexandre Boulch, Renaud Marlet
HIGHLIGHT: Besides, relying on fixed patch sizes may require discretization tuning. To address these issues, we propose to use point cloud convolutions and compute latent vectors at each input point.

202, TITLE: Memory-Augmented Non-Local Attention for Video Super-Resolution
http://openaccess.thecvf.com//content/CVPR2022/html/Yu_Memory-Augmented_Non-Local_Attention_for_Video_Super-Resolution_CVPR_2022_paper.html
AUTHORS: Jiyang Yu, Jingen Liu, Liefeng Bo, Tao Mei
HIGHLIGHT: In this paper, we propose a simple yet effective video super-resolution method that aims at generating high-fidelity high-resolution (HR) videos from low-resolution (LR) ones.

203, TITLE: Neural Texture Extraction and Distribution for Controllable Person Image Synthesis
http://openaccess.thecvf.com//content/CVPR2022/html/Ren_Neural_Texture_Extraction_and_Distribution_for_Controllable_Person_Image_Synthesis_CVPR_2022_paper.html
AUTHORS: Yurui Ren, Xiaoqing Fan, Ge Li, Shan Liu, Thomas H. Li
HIGHLIGHT: Observing that person images are highly structured, we propose to generate desired images by extracting and distributing semantic entities of reference images.

204, TITLE: Classification-Then-Grounding: Reformulating Video Scene Graphs As Temporal Bipartite Graphs
http://openaccess.thecvf.com//content/CVPR2022/html/Gao_Classification-Then-Grounding_Reformulating_Video_Scene_Graphs_As_Temporal_Bipartite_Graphs_CVPR_2022_paper.html
AUTHORS: Kaifeng Gao, Long Chen, Yulei Niu, Jian Shao, Jun Xiao
HIGHLIGHT: Today's VidSGG models are all proposal-based methods, i.e., they first generate numerous paired subject-object snippets as proposals, and then conduct predicate classification for each proposal. In this paper, we argue that this prevalent proposal-based framework has three inherent drawbacks: 1) The ground-truth predicate labels for proposals are partially correct.

205, TITLE: Transformer-Empowered Multi-Scale Contextual Matching and Aggregation for Multi-Contrast MRI Super-Resolution
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Transformer-Empowered_Multi-Scale_Contextual_Matching_and_Aggregation_for_Multi-Contrast_MRI_Super-Resolution_CVPR_2022_paper.html
AUTHORS: Guangyuan Li, Jun Lv, Yapeng Tian, Qi Dou, Chengyan Wang, Chenliang Xu, Jing Qin
HIGHLIGHT: However, existing methods still have two shortcomings: (1) they neglect that the multi-contrast features at different scales contain different anatomical details and hence lack effective mechanisms to match and fuse these features for better reconstruction; and (2) they are still deficient in capturing long-range dependencies, which are essential for the regions with complicated anatomical structures. We propose a novel network to comprehensively address these problems by developing a set of innovative Transformer-empowered multi-scale contextual matching and aggregation techniques; we call it McMRSR.

206, TITLE: GazeOnce: Real-Time Multi-Person Gaze Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_GazeOnce_Real-Time_Multi-Person_Gaze_Estimation_CVPR_2022_paper.html
AUTHORS: Mingfang Zhang, Yunfei Liu, Feng Lu

HIGHLIGHT: In this paper, we propose the first one-stage end-to-end gaze estimation method, GazeOnce, which is capable of simultaneously predicting gaze directions for multiple faces (>10) in an image.

207, TITLE: GateHUB: Gated History Unit With Background Suppression for Online Action Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_GateHUB_Gated_History_Unit_With_Background_Suppression_for_Online_Action_CVPR_2022_paper.html
AUTHORS: Junwen Chen, Gaurav Mittal, Ye Yu, Yu Kong, Mei Chen
HIGHLIGHT: It is therefore important to accentuate parts of the history that are more informative to the prediction of the current frame. We present GateHUB, Gated History Unit with Background Suppression, that comprises a novel position-guided gated cross-attention mechanism to enhance or suppress parts of the history as per how informative they are for current frame prediction.

208, TITLE: Few-Shot Font Generation by Learning Fine-Grained Local Styles
http://openaccess.thecvf.com//content/CVPR2022/html/Tang_Few-Shot_Font_Generation_by_Learning_Fine-Grained_Local_Styles_CVPR_2022_paper.html
AUTHORS: Licheng Tang, Yiyang Cai, Jiaming Liu, Zhibin Hong, Mingming Gong, Minhu Fan, Junyu Han, Jingtuo Liu, Errui Ding, Jingdong Wang
HIGHLIGHT: In this paper, we propose a new font generation approach by learning 1) the fine-grained local styles from references, and 2) the spatial correspondence between the content and reference glyphs.

209, TITLE: Bridging Video-Text Retrieval With Multiple Choice Questions
http://openaccess.thecvf.com//content/CVPR2022/html/Ge_Bridging_Video-Text_Retrieval_With_Multiple_Choice_Questions_CVPR_2022_paper.html
AUTHORS: Yuying Ge, Yixiao Ge, Xihui Liu, Dian Li, Ying Shan, Xiaohu Qie, Ping Luo
HIGHLIGHT: In this work, we enable fine-grained video-text interactions while maintaining high efficiency for retrieval via a novel pretext task, dubbed as Multiple Choice Questions (MCQ), where a parametric module BridgeFormer is trained to answer the "questions" constructed by the text features via resorting to the video features.

210, TITLE: Depth-Aware Generative Adversarial Network for Talking Head Video Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Hong_Depth-Aware_Generative_Adversarial_Network_for_Talking_Head_Video_Generation_CVPR_2022_paper.html
AUTHORS: Fa-Ting Hong, Longhao Zhang, Li Shen, Dan Xu
HIGHLIGHT: In this paper, we introduce a self-supervised face-depth learning method to automatically recover dense 3D facial geometry (i.e. depth) from the face videos without the requirement of any expensive 3D annotation data.

211, TITLE: Dual-Path Image Inpainting With Auxiliary GAN Inversion
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Dual-Path_Image_Inpainting_With_Auxiliary_GAN_Inversion_CVPR_2022_paper.html
AUTHORS: Wentao Wang, Li Niu, Jianfu Zhang, Xue Yang, Liqing Zhang
HIGHLIGHT: In this paper, we develop a dual-path inpainting network with inversion path and feed-forward path, in which inversion path provides auxiliary information to help feed-forward path.

212, TITLE: DF-GAN: A Simple and Effective Baseline for Text-to-Image Synthesis
http://openaccess.thecvf.com//content/CVPR2022/html/Tao_DF-GAN_A_Simple_and_Effective_Baseline_for_Text-to-Image_Synthesis_CVPR_2022_paper.html
AUTHORS: Ming Tao, Hao Tang, Fei Wu, Xiao-Yuan Jing, Bing-Kun Bao, Changsheng Xu
HIGHLIGHT: To these ends, we propose a simpler but more effective Deep Fusion Generative Adversarial Networks (DF-GAN).

213, TITLE: Generative Flows With Invertible Attentions
http://openaccess.thecvf.com//content/CVPR2022/html/Sukthanker_Generative_Flows_With_Invertible_Attentions_CVPR_2022_paper.html
AUTHORS: Rhea Sanjay Sukthanker, Zhiwu Huang, Suryansh Kumar, Radu Timofte, Luc Van Gool
HIGHLIGHT: To fill the gap, this paper introduces two types of invertible attention mechanisms, i.e., map-based and transformer-based attentions, for both unconditional and conditional generative flows.

214, TITLE: Clipped Hyperbolic Classifiers Are Super-Hyperbolic Classifiers
http://openaccess.thecvf.com//content/CVPR2022/html/Guo_Clipped_Hyperbolic_Classifiers_Are_Super-Hyperbolic_Classifiers_CVPR_2022_paper.html
AUTHORS: Yunhui Guo, Xudong Wang, Yubei Chen, Stella X. Yu
HIGHLIGHT: We propose an effective solution by simply clipping the Euclidean feature magnitude while training HNNs.

215, TITLE: Estimating Fine-Grained Noise Model via Contrastive Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Zou_Estimating_Fine-Grained_Noise_Model_via_Contrastive_Learning_CVPR_2022_paper.html
AUTHORS: Yunhao Zou, Ying Fu
HIGHLIGHT: In this work, we combine both noise modeling and estimation, and propose an innovative noise model estimation and noise synthesis pipeline for realistic noisy image generation.

216, TITLE: DiffPoseNet: Direct Differentiable Camera Pose Estimation

23

http://openaccess.thecvf.com//content/CVPR2022/html/Parameshwara_DiffPoseNet_Direct_Differentiable_Camera_Pose_Estimation _CVPR_2022_paper.html
AUTHORS: Chethan M. Parameshwara, Gokul Hari, Cornelia Fermüller, Nitin J. Sanket, Yiannis Aloimonos
HIGHLIGHT: In this paper, we introduce a network NFlowNet, for normal flow estimation which is used to enforce robust and direct constraints.

217, TITLE: The Flag Median and FlagIRLS
http://openaccess.thecvf.com//content/CVPR2022/html/Mankovich_The_Flag_Median_and_FlagIRLS_CVPR_2022_paper.html
AUTHORS: Nathan Mankovich, Emily J. King, Chris Peterson, Michael Kirby
HIGHLIGHT: While a number of different subspace prototypes have been described, the calculation of some of these prototypes has proven to be computationally expensive while other prototypes are affected by outliers and produce highly imperfect clustering on noisy data. This work proposes a new subspace prototype, the flag median, and introduces the FlagIRLS algorithm for its calculation.

218, TITLE: Implicit Feature Decoupling With Depthwise Quantization
http://openaccess.thecvf.com//content/CVPR2022/html/Fostiropoulos_Implicit_Feature_Decoupling_With_Depthwise_Quantization_ CVPR_2022_paper.html
AUTHORS: Iordanis Fostiropoulos, Barry Boehm
HIGHLIGHT: We propose Depthwise Quantization (DQ) where quantization is applied to a decomposed sub-tensor along the feature axis of weak statistical dependence.

219, TITLE: Graph-Context Attention Networks for Size-Varied Deep Graph Matching
http://openaccess.thecvf.com//content/CVPR2022/html/Jiang_Graph-Context_Attention_Networks_for_Size-Varied_Deep_Graph_Matching_CVPR_2022_paper.html
AUTHORS: Zheheng Jiang, Hossein Rahmani, Plamen Angelov, Sue Black, Bryan M. Williams
HIGHLIGHT: To tackle this, we firstly propose to formulate the combinatorial problem of graph matching as an Integer Linear Programming (ILP) problem, which is more flexible and efficient to facilitate comparing graphs of varied sizes. A novel Graph-context Attention Network (GCAN), which jointly capture intrinsic graph structure and cross-graph information for improving the discrimination of node features, is then proposed and trained to resolve this ILP problem with node correspondence supervision.

220, TITLE: FENeRF: Face Editing in Neural Radiance Fields
http://openaccess.thecvf.com//content/CVPR2022/html/Sun_FENeRF_Face_Editing_in_Neural_Radiance_Fields_CVPR_2022_paper .html
AUTHORS: Jingxiang Sun, Xuan Wang, Yong Zhang, Xiaoyu Li, Qi Zhang, Yebin Liu, Jue Wang
HIGHLIGHT: 3D-aware GAN methods can maintain view consistency but their generated images are not locally editable. To overcome these limitations, we propose FENeRF, a 3D-aware generator that can produce view-consistent and locally-editable portrait images.

221, TITLE: CoNeRF: Controllable Neural Radiance Fields
http://openaccess.thecvf.com//content/CVPR2022/html/Kania_CoNeRF_Controllable_Neural_Radiance_Fields_CVPR_2022_paper.h tml
AUTHORS: Kacper Kania, Kwang Moo Yi, Marek Kowalski, Tomasz TrzciÅ„ski, Andrea Tagliasacchi
HIGHLIGHT: Our key idea is to treat the attributes as latent variables that are regressed by the neural network given the scene encoding.

222, TITLE: Noise2NoiseFlow: Realistic Camera Noise Modeling Without Clean Images
http://openaccess.thecvf.com//content/CVPR2022/html/Maleky_Noise2NoiseFlow_Realistic_Camera_Noise_Modeling_Without_Cle an_Images_CVPR_2022_paper.html
AUTHORS: Ali Maleky, Shayan Kousha, Michael S. Brown, Marcus A. Brubaker
HIGHLIGHT: This paper proposes a framework for training a noise model and a denoiser simultaneously while relying only on pairs of noisy images rather than noisy/clean paired image data.

223, TITLE: ZeroWaste Dataset: Towards Deformable Object Segmentation in Cluttered Scenes
http://openaccess.thecvf.com//content/CVPR2022/html/Bashkirova_ZeroWaste_Dataset_Towards_Deformable_Object_Segmentation _in_Cluttered_Scenes_CVPR_2022_paper.html
AUTHORS: Dina Bashkirova, Mohamed Abdelfattah, Ziliang Zhu, James Akl, Fadi Alladkani, Ping Hu, Vitaly Ablavsky, Berk Calli, Sarah Adel Bargal, Kate Saenko
HIGHLIGHT: In this paper, we take a step towards computer-aided waste detection and present the first in-the-wild industrial-grade waste detection and segmentation dataset, ZeroWaste.

224, TITLE: Remember Intentions: Retrospective-Memory-Based Trajectory Prediction
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_Remember_Intentions_Retrospective-Memory-Based_Trajectory_Prediction_CVPR_2022_paper.html
AUTHORS: Chenxin Xu, Weibo Mao, Wenjun Zhang, Siheng Chen
HIGHLIGHT: To provide a more explicit link between the current situation and the seen instances, we imitate the mechanism of retrospective memory in neuropsychology and propose MemoNet, an instance-based approach that predicts the movement intentions of agents by looking for similar scenarios in the training data.

225, TITLE: Measuring Compositional Consistency for Video Question Answering

http://openaccess.thecvf.com//content/CVPR2022/html/Gandhi_Measuring_Compositional_Consistency_for_Video_Question_Answering_CVPR_2022_paper.html
AUTHORS: Mona Gandhi, Mustafa Omer Gul, Eva Prakash, Madeleine Grunde-McLaughlin, Ranjay Krishna, Maneesh Agrawala
HIGHLIGHT: In this paper, we develop a question decomposition engine that programmatically deconstructs a compositional question into a directed acyclic graph of sub-questions.

226, TITLE: Category Contrast for Unsupervised Domain Adaptation in Visual Tasks
http://openaccess.thecvf.com//content/CVPR2022/html/Huang_Category_Contrast_for_Unsupervised_Domain_Adaptation_in_Visual_Tasks_CVPR_2022_paper.html
AUTHORS: Jiaxing Huang, Dayan Guan, Aoran Xiao, Shijian Lu, Ling Shao
HIGHLIGHT: In this work, we explore the idea of instance contrastive learning in unsupervised domain adaptation (UDA) and propose a novel Category Contrast technique (CaCo) that introduces semantic priors on top of instance discrimination for visual UDA tasks.

227, TITLE: SwapMix: Diagnosing and Regularizing the Over-Reliance on Visual Context in Visual Question Answering
http://openaccess.thecvf.com//content/CVPR2022/html/Gupta_SwapMix_Diagnosing_and_Regularizing_the_Over-Reliance_on_Visual_Context_in_CVPR_2022_paper.html
AUTHORS: Vipul Gupta, Zhuowan Li, Adam Kortylewski, Chenyu Zhang, Yingwei Li, Alan Yuille
HIGHLIGHT: In this work, we study the robustness of VQA models from a novel perspective: visual context.

228, TITLE: UNIST: Unpaired Neural Implicit Shape Translation Network
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_UNIST_Unpaired_Neural_Implicit_Shape_Translation_Network_CVPR_2022_paper.html
AUTHORS: Qimin Chen, Johannes Merz, Aditya Sanghi, Hooman Shayani, Ali Mahdavi-Amiri, Hao Zhang
HIGHLIGHT: We introduce UNIST, the first deep neural implicit model for general-purpose, unpaired shape-to-shape translation, in both 2D and 3D domains.

229, TITLE: Local-Adaptive Face Recognition via Graph-Based Meta-Clustering and Regularized Adaptation
http://openaccess.thecvf.com//content/CVPR2022/html/Zhu_Local-Adaptive_Face_Recognition_via_Graph-Based_Meta-Clustering_and_Regularized_Adaptation_CVPR_2022_paper.html
AUTHORS: Wenbin Zhu, Chien-Yi Wang, Kuan-Lun Tseng, Shang-Hong Lai, Baoyuan Wang
HIGHLIGHT: To support continuous learning and fill the last-mile quality gap, we introduce a new problem setup called Local-Adaptive Face Recognition (LaFR).

230, TITLE: The DEVIL Is in the Details: A Diagnostic Evaluation Benchmark for Video Inpainting
http://openaccess.thecvf.com//content/CVPR2022/html/Szeto_The_DEVIL_Is_in_the_Details_A_Diagnostic_Evaluation_Benchmark_CVPR_2022_paper.html
AUTHORS: Ryan Szeto, Jason J. Corso
HIGHLIGHT: Although attributes such as camera and background scene motion inherently change the difficulty of the task and affect methods differently, existing evaluation schemes fail to control for them, thereby providing minimal insight into inpainting failure modes. To address this gap, we propose the Diagnostic Evaluation of Video Inpainting on Landscapes (DEVIL) benchmark, which consists of two contributions: (i) a novel dataset of videos and masks labeled according to several key inpainting failure modes, and (ii) an evaluation scheme that samples slices of the dataset characterized by a fixed content attribute, and scores performance on each slice according to reconstruction, realism, and temporal consistency quality.

231, TITLE: Mutual Information-Driven Pan-Sharpening
http://openaccess.thecvf.com//content/CVPR2022/html/Zhou_Mutual_Information-Driven_Pan-Sharpening_CVPR_2022_paper.html
AUTHORS: Man Zhou, Keyu Yan, Jie Huang, Zihe Yang, Xueyang Fu, Feng Zhao
HIGHLIGHT: This leads to information redundancy not being handled well, which further limits the performance of these methods. To address the above issue, we propose a novel mutual information-driven Pan-sharpening framework in this paper.

232, TITLE: Shifting More Attention to Visual Backbone: Query-Modulated Refinement Networks for End-to-End Visual Grounding
http://openaccess.thecvf.com//content/CVPR2022/html/Ye_Shifting_More_Attention_to_Visual_Backbone_Query-Modulated_Refinement_Networks_for_CVPR_2022_paper.html
AUTHORS: Jiabo Ye, Junfeng Tian, Ming Yan, Xiaoshan Yang, Xuwu Wang, Ji Zhang, Liang He, Xin Lin
HIGHLIGHT: In this paper, we propose a Query-modulated Refinement Network (QRNet) to address the inconsistent issue by adjusting intermediate features in the visual backbone with a novel Query-aware Dynamic Attention (QD-ATT) mechanism and query-aware multiscale fusion.

233, TITLE: A Framework for Learning Ante-Hoc Explainable Models via Concepts
http://openaccess.thecvf.com//content/CVPR2022/html/Sarkar_A_Framework_for_Learning_Ante-Hoc_Explainable_Models_via_Concepts_CVPR_2022_paper.html
AUTHORS: Anirban Sarkar, Deepak Vijaykeerthy, Anindya Sarkar, Vineeth N Balasubramanian
HIGHLIGHT: Self-explaining deep models are designed to learn the latent concept-based explanations implicitly during training, which eliminates the requirement of any post-hoc explanation generation technique. In this work, we propose one such model that appends an explanation generation module on top of any basic network and jointly trains the whole module that shows high predictive performance and generates meaningful explanations in terms of concepts.

234, TITLE: Generating Useful Accident-Prone Driving Scenarios via a Learned Traffic Prior
http://openaccess.thecvf.com/content/CVPR2022/html/Rempe_Generating_Useful_Accident-Prone_Driving_Scenarios_via_a_Learned_Traffic_Prior_CVPR_2022_paper.html
AUTHORS: Davis Rempe, Jonah Philion, Leonidas J. Guibas, Sanja Fidler, Or Litany
HIGHLIGHT: In this work, we introduce STRIVE, a method to automatically generate challenging scenarios that cause a given planner to produce undesirable behavior, like collisions.

235, TITLE: FLOAT: Factorized Learning of Object Attributes for Improved Multi-Object Multi-Part Scene Parsing
http://openaccess.thecvf.com/content/CVPR2022/html/Singh_FLOAT_Factorized_Learning_of_Object_Attributes_for_Improved_Multi-Object_Multi-Part_CVPR_2022_paper.html
AUTHORS: Rishubh Singh, Pranav Gupta, Pradeep Shenoy, Ravikiran Sarvadevabhatla
HIGHLIGHT: In this paper, we propose FLOAT, a factorized label space framework for scalable multi-object multi-part parsing.

236, TITLE: Efficient Geometry-Aware 3D Generative Adversarial Networks
http://openaccess.thecvf.com//content/CVPR2022/html/Chan_Efficient_Geometry-Aware_3D_Generative_Adversarial_Networks_CVPR_2022_paper.html
AUTHORS: Eric R. Chan, Connor Z. Lin, Matthew A. Chan, Koki Nagano, Boxiao Pan, Shalini De Mello, Orazio Gallo, Leonidas J. Guibas, Jonathan Tremblay, Sameh Khamis, Tero Karras, Gordon Wetzstein
HIGHLIGHT: Existing 3D GANs are either compute-intensive or make approximations that are not 3D-consistent; the former limits quality and resolution of the generated images and the latter adversely affects multi-view consistency and shape quality. In this work, we improve the computational efficiency and image quality of 3D GANs without overly relying on these approximations.

237, TITLE: DO-GAN: A Double Oracle Framework for Generative Adversarial Networks
http://openaccess.thecvf.com//content/CVPR2022/html/Aung_DO-GAN_A_Double_Oracle_Framework_for_Generative_Adversarial_Networks_CVPR_2022_paper.html
AUTHORS: Aye Phyu Phyu Aung, Xinrun Wang, Runsheng Yu, Bo An, Senthilnath Jayavelu, Xiaoli Li
HIGHLIGHT: In this paper, we propose a new approach to train Generative Adversarial Networks (GANs) where we deploy a double-oracle framework using the generator and discriminator oracles.

238, TITLE: Dancing Under the Stars: Video Denoising in Starlight
http://openaccess.thecvf.com/content/CVPR2022/html/Monakhova_Dancing_Under_the_Stars_Video_Denoising_in_Starlight_CVPR_2022_paper.html
AUTHORS: Kristina Monakhova, Stephan R. Richter, Laura Waller, Vladlen Koltun
HIGHLIGHT: In this paper, we demonstrate photorealistic video under starlight (no moon present, <0.001 lux) for the first time.

239, TITLE: FocusCut: Diving Into a Focus View in Interactive Segmentation
http://openaccess.thecvf.com/content/CVPR2022/html/Lin_FocusCut_Diving_Into_a_Focus_View_in_Interactive_Segmentation_CVPR_2022_paper.html
AUTHORS: Zheng Lin, Zheng-Peng Duan, Zhao Zhang, Chun-Le Guo, Ming-Ming Cheng
HIGHLIGHT: However, the global view makes the model lose focus from later clicks, and is not in line with user intentions. In this paper, we dive into the view of clicks' eyes to endow them with the decisive role in object details again.

240, TITLE: Medial Spectral Coordinates for 3D Shape Analysis
http://openaccess.thecvf.com/content/CVPR2022/html/Rezanejad_Medial_Spectral_Coordinates_for_3D_Shape_Analysis_CVPR_2022_paper.html
AUTHORS: Morteza Rezanejad, Mohammad Khodadad, Hamidreza Mahyar, Herve Lombaert, Michael Gruninger, Dirk Walther, Kaleem Siddiqi
HIGHLIGHT: Yet, surprisingly, such coordinates have thus far typically considered only local surface positional or derivative information. In the present article, we propose to equip spectral coordinates with medial (object width) information, so as to enrich them.

241, TITLE: Contextualized Spatio-Temporal Contrastive Learning With Self-Supervision
http://openaccess.thecvf.com//content/CVPR2022/html/Yuan_Contextualized_Spatio-Temporal_Contrastive_Learning_With_Self-Supervision_CVPR_2022_paper.html
AUTHORS: Liangzhe Yuan, Rui Qian, Yin Cui, Boqing Gong, Florian Schroff, Ming-Hsuan Yang, Hartwig Adam, Ting Liu
HIGHLIGHT: In this paper, we present Contextualized Spatio-Temporal Contrastive Learning (ConST-CL) to effectively learn spatio-temporally fine-grained video representations via self-supervision.

242, TITLE: Rethinking Architecture Design for Tackling Data Heterogeneity in Federated Learning
http://openaccess.thecvf.com/content/CVPR2022/html/Qu_Rethinking_Architecture_Design_for_Tackling_Data_Heterogeneity_in_Federated_Learning_CVPR_2022_paper.html
AUTHORS: Liangqiong Qu, Yuyin Zhou, Paul Pu Liang, Yingda Xia, Feifei Wang, Ehsan Adeli, Li Fei-Fei, Daniel Rubin
HIGHLIGHT: In this paper, we demonstrate that self-attention-based architectures (e.g., Transformers) are more robust to distribution shifts and hence improve federated learning over heterogeneous data.

243, TITLE: APES: Articulated Part Extraction From Sprite Sheets
http://openaccess.thecvf.com/content/CVPR2022/html/Xu_APES_Articulated_Part_Extraction_From_Sprite_Sheets_CVPR_2022_paper.html

AUTHORS:		Zhan Xu, Matthew Fisher, Yang Zhou, Deepali Aneja, Rushikesh Dudhat, Li Yi, Evangelos Kalogerakis
HIGHLIGHT:		Creating these puppets requires partitioning characters into independently moving parts. In this work, we present a method to automatically identify such articulated parts from a small set of character poses shown in a sprite sheet, which is an illustration of the character that artists often draw before puppet creation.

244, TITLE:		Dressing in the Wild by Watching Dance Videos
http://openaccess.thecvf.com//content/CVPR2022/html/Dong_Dressing_in_the_Wild_by_Watching_Dance_Videos_CVPR_2022_paper.html
AUTHORS:		Xin Dong, Fuwei Zhao, Zhenyu Xie, Xijin Zhang, Daniel K. Du, Min Zheng, Xiang Long, Xiaodan Liang, Jianchao Yang
HIGHLIGHT:		This paper, therefore, attends to virtual try-on in real-world scenes and brings essential improvements in authenticity and naturalness especially for loose garment (e.g., skirts, formal dresses), challenging poses (e.g., cross arms, bent legs), and cluttered backgrounds.

245, TITLE:		SPAct: Self-Supervised Privacy Preservation for Action Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Dave_SPAct_Self-Supervised_Privacy_Preservation_for_Action_Recognition_CVPR_2022_paper.html
AUTHORS:		Ishan Rajendrakumar Dave, Chen Chen, Mubarak Shah
HIGHLIGHT:		For the first time, we present a novel training framework which removes privacy information from input video in a self-supervised manner without requiring privacy labels.

246, TITLE:		Uni6D: A Unified CNN Framework Without Projection Breakdown for 6D Pose Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Jiang_Uni6D_A_Unified_CNN_Framework_Without_Projection_Breakdown_for_6D_CVPR_2022_paper.html
AUTHORS:		Xiaoke Jiang, Donghai Li, Hao Chen, Ye Zheng, Rui Zhao, Liwei Wu
HIGHLIGHT:		As a consequence, the 3D structure is no longer preserved by a modified depth image or feature. To address this issue, we propose a simple yet effective method denoted as Uni6D that explicitly takes the extra UV data along with RGB-D images as input.

247, TITLE:		De-Rendering 3D Objects in the Wild
http://openaccess.thecvf.com//content/CVPR2022/html/Wimbauer_De-Rendering_3D_Objects_in_the_Wild_CVPR_2022_paper.html
AUTHORS:		Felix Wimbauer, Shangzhe Wu, Christian Rupprecht
HIGHLIGHT:		We present a weakly supervised method that is able to decompose a single image of an object into shape (depth and normals), material (albedo, reflectivity and shininess) and global lighting parameters.

248, TITLE:		SPAMs: Structured Implicit Parametric Models
http://openaccess.thecvf.com//content/CVPR2022/html/Palafox_SPAMs_Structured_Implicit_Parametric_Models_CVPR_2022_paper.html
AUTHORS:		Pablo Palafox, Nikolaos Sarafianos, Tony Tung, Angela Dai
HIGHLIGHT:		We observe that deformable object motion is often semantically structured, and thus propose to learn Structured-implicit PArametric Models (SPAMs) as a deformable object representation that structurally decomposes non-rigid object motion into part-based disentangled representations of shape and pose, with each being represented by deep implicit functions.

249, TITLE:		Global Sensing and Measurements Reuse for Image Compressed Sensing
http://openaccess.thecvf.com//content/CVPR2022/html/Fan_Global_Sensing_and_Measurements_Reuse_for_Image_Compressed_Sensing_CVPR_2022_paper.html
AUTHORS:		Zi-En Fan, Feng Lian, Jia-Ni Quan
HIGHLIGHT:		Moreover, using measurements only once may not be enough for extracting richer information from measurements. To address these issues, we propose a novel Measurements Reuse Convolutional Compressed Sensing Network (MR-CCSNet) which employs Global Sensing Module (GSM) to collect all level features for achieving an efficient sensing and Measurements Reuse Block (MRB) to reuse measurements multiple times on multi-scale.

250, TITLE:		SeeThroughNet: Resurrection of Auxiliary Loss by Preserving Class Probability Information
http://openaccess.thecvf.com//content/CVPR2022/html/Han_SeeThroughNet_Resurrection_of_Auxiliary_Loss_by_Preserving_Class_Probability_Information_CVPR_2022_paper.html
AUTHORS:		Dasol Han, Jaewook Yoo, Dokwan Oh
HIGHLIGHT:		In this paper, we introduce Class Probability Preserving (CPP) pooling to alleviate information loss in down-sampling the ground truth in semantic segmentation tasks.

251, TITLE:		Representing 3D Shapes With Probabilistic Directed Distance Fields
http://openaccess.thecvf.com//content/CVPR2022/html/Aumentado-Armstrong_Representing_3D_Shapes_With_Probabilistic_Directed_Distance_Fields_CVPR_2022_paper.html
AUTHORS:		Tristan Aumentado-Armstrong, Stavros Tsogkas, Sven Dickinson, Allan D. Jepson
HIGHLIGHT:		In this work, we endeavour to address both shortcomings with a novel shape representation that allows fast differentiable rendering within an implicit architecture.

252, TITLE:		Learning ABCs: Approximate Bijective Correspondence for Isolating Factors of Variation With Weak Supervision
http://openaccess.thecvf.com//content/CVPR2022/html/Murphy_Learning_ABCs_Approximate_Bijective_Correspondence_for_Isolating_Factors_of_Variation_CVPR_2022_paper.html

AUTHORS: Kieran A. Murphy, Varun Jampani, Srikumar Ramalingam, Ameesh Makadia
HIGHLIGHT: We propose a novel algorithm that utilizes a weak form of supervision where the data is partitioned into sets according to certain inactive (common) factors of variation which are invariant across elements of each set.

253, TITLE: ABO: Dataset and Benchmarks for Real-World 3D Object Understanding
http://openaccess.thecvf.com//content/CVPR2022/html/Collins_ABO_Dataset_and_Benchmarks_for_Real-World_3D_Object_Understanding_CVPR_2022_paper.html
AUTHORS: Jasmine Collins, Shubham Goel, Kenan Deng, Achleshwar Luthra, Leon Xu, Erhan Gundogdu, Xi Zhang, Tomas F. Yago Vicente, Thomas Dideriksen, Himanshu Arora, Matthieu Guillaumin, Jitendra Malik
HIGHLIGHT: We introduce Amazon Berkeley Objects (ABO), a new large-scale dataset designed to help bridge the gap between real and virtual 3D worlds.

254, TITLE: DETReg: Unsupervised Pretraining With Region Priors for Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Bar_DETReg_Unsupervised_Pretraining_With_Region_Priors_for_Object_Detection_CVPR_2022_paper.html
AUTHORS: Amir Bar, Xin Wang, Vadim Kantorov, Colorado J. Reed, Roei Herzig, Gal Chechik, Anna Rohrbach, Trevor Darrell, Amir Globerson
HIGHLIGHT: Instead, we introduce DETReg, a new self-supervised method that pretrains the entire object detection network, including the object localization and embedding components.

255, TITLE: Learning To Restore 3D Face From In-the-Wild Degraded Images
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Learning_To_Restore_3D_Face_From_In-the-Wild_Degraded_Images_CVPR_2022_paper.html
AUTHORS: Zhenyu Zhang, Yanhao Ge, Ying Tai, Xiaoming Huang, Chengjie Wang, Hao Tang, Dongjin Huang, Zhifeng Xie
HIGHLIGHT: In-the-wild 3D face modelling is a challenging problem as the predicted facial geometry and texture suffer from a lack of reliable clues or priors, when the input images are degraded. To address such a problem, in this paper we propose a novel Learning to Restore (L2R) 3D face framework for unsupervised high-quality face reconstruction from low-resolution images.

256, TITLE: Practical Evaluation of Adversarial Robustness via Adaptive Auto Attack
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Practical_Evaluation_of_Adversarial_Robustness_via_Adaptive_Auto_Attack_CVPR_2022_paper.html
AUTHORS: Ye Liu, Yaya Cheng, Lianli Gao, Xianglong Liu, Qilong Zhang, Jingkuan Song
HIGHLIGHT: A practical evaluation method should be convenient (i.e., parameter-free), efficient (i.e., fewer iterations) and reliable (i.e., approaching the lower bound of robustness). Towards this target, we propose a parameter-free Adaptive Auto Attack (A3) evaluation method which addresses the efficiency and reliability in a test-time-training fashion.

257, TITLE: Convolutions for Spatial Interaction Modeling
http://openaccess.thecvf.com//content/CVPR2022/html/Su_Convolutions_for_Spatial_Interaction_Modeling_CVPR_2022_paper.html
AUTHORS: Zhaoen Su, Chao Wang, David Bradley, Carlos Vallespi-Gonzalez, Carl Wellington, Nemanja Djuric
HIGHLIGHT: In this paper, we consider the problem of spatial interaction modeling in the context of predicting the motion of actors around autonomous vehicles, and investigate alternatives to GNNs.

258, TITLE: MS-TCT: Multi-Scale Temporal ConvTransformer for Action Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Dai_MS-TCT_Multi-Scale_Temporal_ConvTransformer_for_Action_Detection_CVPR_2022_paper.html
AUTHORS: Rui Dai, Srijan Das, Kumara Kahatapitiya, Michael S. Ryoo, François Brémond
HIGHLIGHT: For detecting actions in those complex videos, efficiently capturing both short-term and long-term temporal information in the video is critical. To this end, we propose a novel ConvTransformer network for action detection.

259, TITLE: Salvage of Supervision in Weakly Supervised Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Sui_Salvage_of_Supervision_in_Weakly_Supervised_Object_Detection_CVPR_2022_paper.html
AUTHORS: Lin Sui, Chen-Lin Zhang, Jianxin Wu
HIGHLIGHT: To bridge the performance and technical gaps between WSOD and FSOD, this paper proposes a new framework, Salvage of Supervision (SoS), with the key idea being to harness every potentially useful supervisory signal in WSOD: the weak image-level labels, the pseudo-labels, and the power of semi-supervised object detection.

260, TITLE: Cross-View Transformers for Real-Time Map-View Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Zhou_Cross-View_Transformers_for_Real-Time_Map-View_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Brady Zhou, Philipp Kräuhenbühl
HIGHLIGHT: We present cross-view transformers, an efficient attention-based model for map-view semantic segmentation from multiple cameras.

261, TITLE: Distinguishing Unseen From Seen for Generalized Zero-Shot Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Su_Distinguishing_Unseen_From_Seen_for_Generalized_Zero-Shot_Learning_CVPR_2022_paper.html
AUTHORS: Hongzu Su, Jingjing Li, Zhi Chen, Lei Zhu, Ke Lu

HIGHLIGHT: In this paper, we present a novel method which leverages both visual and semantic modalities to distinguish seen and unseen categories.

262, TITLE: Online Continual Learning on a Contaminated Data Stream With Blurry Task Boundaries
http://openaccess.thecvf.com//content/CVPR2022/html/Bang_Online_Continual_Learning_on_a_Contaminated_Data_Stream_With_Blurry_CVPR_2022_paper.html
AUTHORS: Jihwan Bang, Hyunseo Koh, Seulki Park, Hwanjun Song, Jung-Woo Ha, Jonghyun Choi
HIGHLIGHT: To balance diversity and purity in the episodic memory, we propose a novel strategy to manage and use the memory by a unified approach of label noise aware diverse sampling and robust learning with semi-supervised learning.

263, TITLE: Controllable Dynamic Multi-Task Architectures
http://openaccess.thecvf.com//content/CVPR2022/html/Raychaudhuri_Controllable_Dynamic_Multi-Task_Architectures_CVPR_2022_paper.html
AUTHORS: Dripta S. Raychaudhuri, Yumin Suh, Samuel Schulter, Xiang Yu, Masoud Faraki, Amit K. Roy-Chowdhury, Manmohan Chandraker
HIGHLIGHT: This challenge motivates models which allow control over the relative importance of tasks and total compute cost during inference time. In this work, we propose such a controllable multi-task network that dynamically adjusts its architecture and weights to match the desired task preference as well as the resource constraints.

264, TITLE: Learning To Imagine: Diversify Memory for Incremental Learning Using Unlabeled Data
http://openaccess.thecvf.com//content/CVPR2022/html/Tang_Learning_To_Imagine_Diversify_Memory_for_Incremental_Learning_Using_Unlabeled_CVPR_2022_paper.html
AUTHORS: Yu-Ming Tang, Yi-Xing Peng, Wei-Shi Zheng
HIGHLIGHT: Although maintaining a handful of samples (called "exemplars") of each task could alleviate forgetting to some extent, existing methods are still limited by the small number of exemplars since these exemplars are too few to carry enough task-specific knowledge, and therefore the forgetting remains. To overcome this problem, we propose to "imagine" diverse counterparts of given exemplars referring to the abundant semantic-irrelevant information from unlabeled data.

265, TITLE: SmartAdapt: Multi-Branch Object Detection Framework for Videos on Mobiles
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_SmartAdapt_Multi-Branch_Object_Detection_Framework_for_Videos_on_Mobiles_CVPR_2022_paper.html
AUTHORS: Ran Xu, Fangzhou Mu, Jayoung Lee, Preeti Mukherjee, Somali Chaterji, Saurabh Bagchi, Yin Li
HIGHLIGHT: In this paper, we ask, and answer, the wide-ranging question across all MBODFs: How to expose the right set of execution branches and then how to schedule the optimal one at inference time?

266, TITLE: VL-Adapter: Parameter-Efficient Transfer Learning for Vision-and-Language Tasks
http://openaccess.thecvf.com//content/CVPR2022/html/Sung_VL-Adapter_Parameter-Efficient_Transfer_Learning_for_Vision-and-Language_Tasks_CVPR_2022_paper.html
AUTHORS: Yi-Lin Sung, Jaemin Cho, Mohit Bansal
HIGHLIGHT: Hence, in this paper, we introduce adapter-based parameter-efficient transfer learning techniques to V&L models such as VL-BART and VL-T5.

267, TITLE: Deep Hybrid Models for Out-of-Distribution Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Cao_Deep_Hybrid_Models_for_Out-of-Distribution_Detection_CVPR_2022_paper.html
AUTHORS: Senqi Cao, Zhongfei Zhang
HIGHLIGHT: We propose a principled and practical method for out-of-distribution (OoD) detection with deep hybrid models (DHMs), which model the joint density $p(x,y)$ of features and labels with a single forward pass.

268, TITLE: Accelerating Video Object Segmentation With Compressed Video
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_Accelerating_Video_Object_Segmentation_With_Compressed_Video_CVPR_2022_paper.html
AUTHORS: Kai Xu, Angela Yao
HIGHLIGHT: We propose an efficient plug-and-play acceleration framework for semi-supervised video object segmentation by exploiting the temporal redundancies in videos presented by the compressed bitstream.

269, TITLE: Exploring Domain-Invariant Parameters for Source Free Domain Adaptation
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Exploring_Domain-Invariant_Parameters_for_Source_Free_Domain_Adaptation_CVPR_2022_paper.html
AUTHORS: Fan Wang, Zhongyi Han, Yongshun Gong, Yilong Yin
HIGHLIGHT: The motivation behind this insight is clear: the domain-invariant representations are dominated by only partial parameters of an available deep source model. We devise the Domain-Invariant Parameter Exploring (DIPE) approach to capture such domain-invariant parameters in the source model to generate domain-invariant representations.

270, TITLE: FastDOG: Fast Discrete Optimization on GPU
http://openaccess.thecvf.com//content/CVPR2022/html/Abbas_FastDOG_Fast_Discrete_Optimization_on_GPU_CVPR_2022_paper.html
AUTHORS: Ahmed Abbas, Paul Swoboda

29

HIGHLIGHT: We present a massively parallel Lagrange decomposition method for solving 0--1 integer linear programs occurring in structured prediction.

271, TITLE: Fire Together Wire Together: A Dynamic Pruning Approach With Self-Supervised Mask Prediction
http://openaccess.thecvf.com//content/CVPR2022/html/Elkerdawy_Fire_Together_Wire_Together_A_Dynamic_Pruning_Approach_With_Self-Supervised_CVPR_2022_paper.html
AUTHORS: Sara Elkerdawy, Mostafa Elhoushi, Hong Zhang, Nilanjan Ray
HIGHLIGHT: Our contribution is two-fold: 1) decoupled task and pruning training. 2) Simple hyperparameter selection that enables FLOPs reduction estimation before training.

272, TITLE: Multi-Source Uncertainty Mining for Deep Unsupervised Saliency Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Multi-Source_Uncertainty_Mining_for_Deep_Unsupervised_Saliency_Detection_CVPR_2022_paper.html
AUTHORS: Yifan Wang, Wenbo Zhang, Lijun Wang, Ting Liu, Huchuan Lu
HIGHLIGHT: In this paper, we propose a novel multi-source uncertainty mining method to facilitate unsupervised deep learning from multiple noisy labels generated by traditional handcrafted SOD methods.

273, TITLE: Self-Supervised Equivariant Learning for Oriented Keypoint Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Lee_Self-Supervised_Equivariant_Learning_for_Oriented_Keypoint_Detection_CVPR_2022_paper.html
AUTHORS: Jongmin Lee, Byungjin Kim, Minsu Cho
HIGHLIGHT: To learn to detect robust oriented keypoints, we introduce a self-supervised learning framework using rotation-equivariant CNNs.

274, TITLE: Wavelet Knowledge Distillation: Towards Efficient Image-to-Image Translation
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Wavelet_Knowledge_Distillation_Towards_Efficient_Image-to-Image_Translation_CVPR_2022_paper.html
AUTHORS: Linfeng Zhang, Xin Chen, Xiaobing Tu, Pengfei Wan, Ning Xu, Kaisheng Ma
HIGHLIGHT: The results show that GANs, especially small GANs lack the ability to generate high-quality high frequency information. To address this problem, we propose a novel knowledge distillation method referred to as wavelet knowledge distillation.

275, TITLE: Focal and Global Knowledge Distillation for Detectors
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Focal_and_Global_Knowledge_Distillation_for_Detectors_CVPR_2022_paper.html
AUTHORS: Zhendong Yang, Zhe Li, Xiaohu Jiang, Yuan Gong, Zehuan Yuan, Danpei Zhao, Chun Yuan
HIGHLIGHT: In this paper, we point out that in object detection, the features of the teacher and student vary greatly in different areas, especially in the foreground and background.

276, TITLE: Learning To Prompt for Continual Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Learning_To_Prompt_for_Continual_Learning_CVPR_2022_paper.html
AUTHORS: Zifeng Wang, Zizhao Zhang, Chen-Yu Lee, Han Zhang, Ruoxi Sun, Xiaoqi Ren, Guolong Su, Vincent Perot, Jennifer Dy, Tomas Pfister
HIGHLIGHT: Typical methods rely on a rehearsal buffer or known task identity at test time to retrieve learned knowledge and address forgetting, while this work presents a new paradigm for continual learning that aims to train a more succinct memory system without accessing task identity at test time.

277, TITLE: Human Mesh Recovery From Multiple Shots
http://openaccess.thecvf.com//content/CVPR2022/html/Pavlakos_Human_Mesh_Recovery_From_Multiple_Shots_CVPR_2022_paper.html
AUTHORS: Georgios Pavlakos, Jitendra Malik, Angjoo Kanazawa
HIGHLIGHT: However, the richness of data comes at the expense of fundamental challenges such as abrupt shot changes and close up shots of actors with heavy truncation, which limits the applicability of existing 3D human understanding methods. In this paper, we address these limitations with the insight that while shot changes of the same scene incur a discontinuity between frames, the 3D structure of the scene still changes smoothly.

278, TITLE: Improving Adversarial Transferability via Neuron Attribution-Based Attacks
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Improving_Adversarial_Transferability_via_Neuron_Attribution-Based_Attacks_CVPR_2022_paper.html
AUTHORS: Jianping Zhang, Weibin Wu, Jen-tse Huang, Yizhan Huang, Wenxuan Wang, Yuxin Su, Michael R. Lyu
HIGHLIGHT: However, existing feature-level attacks generally employ inaccurate neuron importance estimations, which deteriorates their transferability. To overcome such pitfalls, in this paper, we propose the Neuron Attribution-based Attack (NAA), which conducts feature-level attacks with more accurate neuron importance estimations.

279, TITLE: Better Trigger Inversion Optimization in Backdoor Scanning
http://openaccess.thecvf.com//content/CVPR2022/html/Tao_Better_Trigger_Inversion_Optimization_in_Backdoor_Scanning_CVPR_2022_paper.html
AUTHORS: Guanhong Tao, Guangyu Shen, Yingqi Liu, Shengwei An, Qiuling Xu, Shiqing Ma, Pan Li, Xiangyu Zhang
HIGHLIGHT: We develop a new optimization method that directly minimizes individual pixel changes, without using a mask.

280, TITLE: GANSeg: Learning To Segment by Unsupervised Hierarchical Image Generation
http://openaccess.thecvf.com//content/CVPR2022/html/He_GANSeg_Learning_To_Segment_by_Unsupervised_Hierarchical_Image_Generation_CVPR_2022_paper.html
AUTHORS: Xingzhe He, Bastian Wandt, Helge Rhodin
HIGHLIGHT: Weakly-supervised and unsupervised methods exist, but they depend on the comparison of pairs of images, such as from multi-views, frames of videos, and image augmentation, which limits their applicability. To address this, we propose a GAN-based approach that generates images conditioned on latent masks, thereby alleviating full or weak annotations required in previous approaches.

281, TITLE: Dense Learning Based Semi-Supervised Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Dense_Learning_Based_Semi-Supervised_Object_Detection_CVPR_2022_paper.html
AUTHORS: Binghui Chen, Pengyu Li, Xiang Chen, Biao Wang, Lei Zhang, Xian-Sheng Hua
HIGHLIGHT: Although a few works have proposed various self-training-based methods or consistency-regularization-based methods, they all target anchor-based detectors, while ignoring the dependency on anchor-free detectors of the actual industrial deployment. To this end, in this paper, we intend to bridge the gap on anchor-free SSOD algorithm by proposing a DenSe Learning (DSL) based algorithm for SSOD.

282, TITLE: Fixing Malfunctional Objects With Learned Physical Simulation and Functional Prediction
http://openaccess.thecvf.com//content/CVPR2022/html/Hong_Fixing_Malfunctional_Objects_With_Learned_Physical_Simulation_and_Functional_Prediction_CVPR_2022_paper.html
AUTHORS: Yining Hong, Kaichun Mo, Li Yi, Leonidas J. Guibas, Antonio Torralba, Joshua B. Tenenbaum, Chuang Gan
HIGHLIGHT: To mimic humans' mental simulation process, we present FixNet, a novel framework that seamlessly incorporates perception and physical dynamics.

283, TITLE: Convolution of Convolution: Let Kernels Spatially Collaborate
http://openaccess.thecvf.com//content/CVPR2022/html/Zhao_Convolution_of_Convolution_Let_Kernels_Spatially_Collaborate_CVPR_2022_paper.html
AUTHORS: Rongzhen Zhao, Jian Li, Zhenzhi Wu
HIGHLIGHT: In the biological visual pathway, especially the retina, neurons are tiled along spatial dimensions with the electrical coupling as their local association, while in a convolution layer, kernels are placed along the channel dimension singly. We propose Convolution of Convolution, associating kernels in a layer and letting them collaborate spatially.

284, TITLE: Make It Move: Controllable Image-to-Video Generation With Text Descriptions
http://openaccess.thecvf.com//content/CVPR2022/html/Hu_Make_It_Move_Controllable_Image-to-Video_Generation_With_Text_Descriptions_CVPR_2022_paper.html
AUTHORS: Yaosi Hu, Chong Luo, Zhenzhong Chen
HIGHLIGHT: The key challenges of TI2V task lie both in aligning appearance and motion from different modalities, and in handling uncertainty in text descriptions. To address these challenges, we propose a Motion Anchor-based video GEnerator (MAGE) with an innovative motion anchor (MA) structure to store appearance-motion aligned representation.

285, TITLE: C2AM Loss: Chasing a Better Decision Boundary for Long-Tail Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_C2AM_Loss_Chasing_a_Better_Decision_Boundary_for_Long-Tail_Object_CVPR_2022_paper.html
AUTHORS: Tong Wang, Yousong Zhu, Yingying Chen, Chaoyang Zhao, Bin Yu, Jinqiao Wang, Ming Tang
HIGHLIGHT: Hence, we devise a Category-Aware Angular Margin Loss (C2AM Loss) to introduce an adaptive angular margin between any two categories

286, TITLE: Neural Points: Point Cloud Representation With Neural Fields for Arbitrary Upsampling
http://openaccess.thecvf.com//content/CVPR2022/html/Feng_Neural_Points_Point_Cloud_Representation_With_Neural_Fields_for_Arbitrary_CVPR_2022_paper.html
AUTHORS: Wanquan Feng, Jin Li, Hongrui Cai, Xiaonan Luo, Juyong Zhang
HIGHLIGHT: In this paper, we propose Neural Points, a novel point cloud representation and apply it to the arbitrary-factored upsampling task.

287, TITLE: Distribution Consistent Neural Architecture Search
http://openaccess.thecvf.com//content/CVPR2022/html/Pan_Distribution_Consistent_Neural_Architecture_Search_CVPR_2022_paper.html
AUTHORS: Junyi Pan, Chong Sun, Yizhou Zhou, Ying Zhang, Chen Li
HIGHLIGHT: Instead, in this paper, we propose a novel distribution consistent one-shot neural architecture search algorithm.

288, TITLE: Video-Text Representation Learning via Differentiable Weak Temporal Alignment
http://openaccess.thecvf.com//content/CVPR2022/html/Ko_Video-Text_Representation_Learning_via_Differentiable_Weak_Temporal_Alignment_CVPR_2022_paper.html
AUTHORS: Dohwan Ko, Joonmyung Choi, Juyeon Ko, Shinyeong Noh, Kyoung-Woon On, Eun-Sol Kim, Hyunwoo J. Kim
HIGHLIGHT: In this paper, we propose a novel multi-modal self-supervised framework Video-Text Temporally Weak Alignment-based Contrastive Learning (VT-TWINS) to capture significant information from noisy and weakly correlated data using a variant of Dynamic Time Warping (DTW).

289, TITLE: Bi-Directional Object-Context Prioritization Learning for Saliency Ranking

http://openaccess.thecvf.com//content/CVPR2022/html/Tian_Bi-Directional_Object-Context_Prioritization_Learning_for_Saliency_Ranking_CVPR_2022_paper.html
AUTHORS: Xin Tian, Ke Xu, Xin Yang, Lin Du, Baocai Yin, Rynson W.H. Lau
HIGHLIGHT: This inspires us to model the region-level interactions, in addition to the object-level reasoning, for saliency ranking. To this end, we propose a novel bi-directional method to unify spatial attention and object-based attention for saliency ranking.

290, TITLE: FreeSOLO: Learning To Segment Objects Without Annotations
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_FreeSOLO_Learning_To_Segment_Objects_Without_Annotations_CVPR_2022_paper.html
AUTHORS: Xinlong Wang, Zhiding Yu, Shalini De Mello, Jan Kautz, Anima Anandkumar, Chunhua Shen, Jose M. Alvarez
HIGHLIGHT: We present FreeSOLO, a self-supervised instance segmentation framework built on top of the simple instance segmentation method SOLO.

291, TITLE: What Do Navigation Agents Learn About Their Environment?
http://openaccess.thecvf.com//content/CVPR2022/html/Dwivedi_What_Do_Navigation_Agents_Learn_About_Their_Environment_CVPR_2022_paper.html
AUTHORS: Kshitij Dwivedi, Gemma Roig, Aniruddha Kembhavi, Roozbeh Mottaghi
HIGHLIGHT: In this paper, we introduce the Interpretability System for Embodied agEnts (iSEE) for Point Goal (PointNav) and Object Goal (ObjectNav) navigation models.

292, TITLE: Progressive Minimal Path Method With Embedded CNN
http://openaccess.thecvf.com//content/CVPR2022/html/Liao_Progressive_Minimal_Path_Method_With_Embedded_CNN_CVPR_2022_paper.html
AUTHORS: Wei Liao
HIGHLIGHT: We propose Path-CNN, a method for the segmentation of centerlines of tubular structures by embedding convolutional neural networks (CNNs) into the progressive minimal path method.

293, TITLE: FIFO: Learning Fog-Invariant Features for Foggy Scene Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Lee_FIFO_Learning_Fog-Invariant_Features_for_Foggy_Scene_Segmentation_CVPR_2022_paper.html
AUTHORS: Sohyun Lee, Taeyoung Son, Suha Kwak
HIGHLIGHT: Robust visual recognition under adverse weather conditions is of great importance in real-world applications. In this context, we propose a new method for learning semantic segmentation models robust against fog.

294, TITLE: 3D Human Tongue Reconstruction From Single "In-the-Wild" Images
http://openaccess.thecvf.com//content/CVPR2022/html/Ploumpis_3D_Human_Tongue_Reconstruction_From_Single_In-the-Wild_Images_CVPR_2022_paper.html
AUTHORS: Stylianos Ploumpis, Stylianos Moschoglou, Vasileios Triantafyllou, Stefanos Zafeiriou
HIGHLIGHT: In this work we present the first, to the best of our knowledge, end-to-end trainable pipeline that accurately reconstructs the 3D face together with the tongue.

295, TITLE: Enhancing Adversarial Robustness for Deep Metric Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Zhou_Enhancing_Adversarial_Robustness_for_Deep_Metric_Learning_CVPR_2022_paper.html
AUTHORS: Mo Zhou, Vishal M. Patel
HIGHLIGHT: Conversely, we propose Hardness Manipulation to efficiently perturb the training triplet till a specified level of hardness for adversarial training, according to a harder benign triplet or a pseudo-hardness function.

296, TITLE: Multi-Scale High-Resolution Vision Transformer for Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Gu_Multi-Scale_High-Resolution_Vision_Transformer_for_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Jiaqi Gu, Hyoukjun Kwon, Dilin Wang, Wei Ye, Meng Li, Yu-Hsin Chen, Liangzhen Lai, Vikas Chandra, David Z. Pan
HIGHLIGHT: However, ViTs are mainly designed for image classification that generate single-scale low-resolution representations, which makes dense prediction tasks such as semantic segmentation challenging for ViTs. Therefore, we propose HRViT, which enhances ViTs to learn semantically-rich and spatially-precise multi-scale representations by integrating high-resolution multi-branch architectures with ViTs.

297, TITLE: Lite-MDETR: A Lightweight Multi-Modal Detector
http://openaccess.thecvf.com//content/CVPR2022/html/Lou_Lite-MDETR_A_Lightweight_Multi-Modal_Detector_CVPR_2022_paper.html
AUTHORS: Qian Lou, Yen-Chang Hsu, Burak Uzkent, Ting Hua, Yilin Shen, Hongxia Jin
HIGHLIGHT: In this paper, we present a Lightweight modulated detector, Lite-MDETR, to facilitate efficient end-to-end multi-modal understanding on mobile devices.

298, TITLE: CoordGAN: Self-Supervised Dense Correspondences Emerge From GANs
http://openaccess.thecvf.com//content/CVPR2022/html/Mu_CoordGAN_Self-Supervised_Dense_Correspondences_Emerge_From_GANs_CVPR_2022_paper.html
AUTHORS: Jiteng Mu, Shalini De Mello, Zhiding Yu, Nuno Vasconcelos, Xiaolong Wang, Jan Kautz, Sifei Liu

HIGHLIGHT: In this work, we introduce Coordinate GAN (CoordGAN), a structure-texture disentangled GAN that learns a dense correspondence map for each generated image.

299, TITLE: A Simple Multi-Modality Transfer Learning Baseline for Sign Language Translation
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_A_Simple_Multi-Modality_Transfer_Learning_Baseline_for_Sign_Language_Translation_CVPR_2022_paper.html
AUTHORS: Yutong Chen, Fangyun Wei, Xiao Sun, Zhirong Wu, Stephen Lin
HIGHLIGHT: This paper proposes a simple transfer learning baseline for sign language translation.

300, TITLE: Unsupervised Visual Representation Learning by Online Constrained K-Means
http://openaccess.thecvf.com//content/CVPR2022/html/Qian_Unsupervised_Visual_Representation_Learning_by_Online_Constrained_K-Means_CVPR_2022_paper.html
AUTHORS: Qi Qian, Yuanhong Xu, Juhua Hu, Hao Li, Rong Jin
HIGHLIGHT: To address these challenges, we first investigate the objective of clustering-based representation learning. Based on this, we propose a novel clustering-based pretext task with online Constrained K-means (CoKe).

301, TITLE: Neural Point Light Fields
http://openaccess.thecvf.com//content/CVPR2022/html/Ost_Neural_Point_Light_Fields_CVPR_2022_paper.html
AUTHORS: Julian Ost, Issam Laradji, Alejandro Newell, Yuval Bahat, Felix Heide
HIGHLIGHT: We introduce Neural Point Light Fields that represent scenes implicitly with a light field living on a sparse point cloud.

302, TITLE: Vehicle Trajectory Prediction Works, but Not Everywhere
http://openaccess.thecvf.com//content/CVPR2022/html/Bahari_Vehicle_Trajectory_Prediction_Works_but_Not_Everywhere_CVPR_2022_paper.html
AUTHORS: Mohammadhossein Bahari, Saeed Saadatnejad, Ahmad Rahimi, Mohammad Shaverdikondori, Amir Hossein Shahidzadeh, Seyed-Mohsen Moosavi-Dezfooli, Alexandre Alahi
HIGHLIGHT: We present a novel method that automatically generates realistic scenes causing state-of-the-art models to go off-road.

303, TITLE: PSMNet: Position-Aware Stereo Merging Network for Room Layout Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_PSMNet_Position-Aware_Stereo_Merging_Network_for_Room_Layout_Estimation_CVPR_2022_paper.html
AUTHORS: Haiyan Wang, Will Hutchcroft, Yuguang Li, Zhiqiang Wan, Ivaylo Boyadzhiev, Yingli Tian, Sing Bing Kang
HIGHLIGHT: In this paper, we propose a new deep learning-based method for estimating room layout given a pair of 360 panoramas.

304, TITLE: MonoDTR: Monocular 3D Object Detection With Depth-Aware Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Huang_MonoDTR_Monocular_3D_Object_Detection_With_Depth-Aware_Transformer_CVPR_2022_paper.html
AUTHORS: Kuan-Chih Huang, Tsung-Han Wu, Hung-Ting Su, Winston H. Hsu
HIGHLIGHT: Some existing methods leverage depth information from an off-the-shelf depth estimator to assist 3D detection, but suffer from the additional computational burden and achieve limited performance caused by inaccurate depth priors. To alleviate this, we propose MonoDTR, a novel end-to-end depth-aware transformer network for monocular 3D object detection.

305, TITLE: Learning Graph Regularisation for Guided Super-Resolution
http://openaccess.thecvf.com//content/CVPR2022/html/de_Lutio_Learning_Graph_Regularisation_for_Guided_Super-Resolution_CVPR_2022_paper.html
AUTHORS: Riccardo de Lutio, Alexander Becker, Stefano D'Aronco, Stefania Russo, Jan D. Wegner, Konrad Schindler
HIGHLIGHT: We introduce a novel formulation for guided super-resolution.

306, TITLE: Instance-Wise Occlusion and Depth Orders in Natural Scenes
http://openaccess.thecvf.com//content/CVPR2022/html/Lee_Instance-Wise_Occlusion_and_Depth_Orders_in_Natural_Scenes_CVPR_2022_paper.html
AUTHORS: Hyunmin Lee, Jaesik Park
HIGHLIGHT: In this paper, we introduce a new dataset, named InstaOrder, that can be used to understand the spatial relationships of instances in a 3D space.

307, TITLE: Look for the Change: Learning Object States and State-Modifying Actions From Untrimmed Web Videos
http://openaccess.thecvf.com//content/CVPR2022/html/Soucek_Look_for_the_Change_Learning_Object_States_and_State-Modifying_Actions_CVPR_2022_paper.html
AUTHORS: Tomáš Souček, Jean-Baptiste Alayrac, Antoine Miech, Ivan Laptev, Josef Sivic
HIGHLIGHT: In this paper, we seek to temporally localize object states (e.g. "empty" and "full" cup) together with the corresponding state-modifying actions ("pouring coffee") in long uncurated videos with minimal supervision.

308, TITLE: Attribute Surrogates Learning and Spectral Tokens Pooling in Transformers for Few-Shot Learning
http://openaccess.thecvf.com//content/CVPR2022/html/He_Attribute_Surrogates_Learning_and_Spectral_Tokens_Pooling_in_Transformers_for_CVPR_2022_paper.html
AUTHORS: Yangji He, Weihan Liang, Dongyang Zhao, Hong-Yu Zhou, Weifeng Ge, Yizhou Yu, Wenqiang Zhang

HIGHLIGHT: This paper presents new hierarchically cascaded transformers that can improve data efficiency through attribute surrogates learning and spectral tokens pooling.

309, TITLE: Generalized Category Discovery
http://openaccess.thecvf.com//content/CVPR2022/html/Vaze_Generalized_Category_Discovery_CVPR_2022_paper.html
AUTHORS: Sagar Vaze, Kai Han, Andrea Vedaldi, Andrew Zisserman
HIGHLIGHT: In this paper, we consider a highly general image recognition setting wherein, given a labelled and unlabelled set of images, the task is to categorize all images in the unlabelled set.

310, TITLE: Maximum Consensus by Weighted Influences of Monotone Boolean Functions
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Maximum_Consensus_by_Weighted_Influences_of_Monotone_Boolean_Functions_CVPR_2022_paper.html
AUTHORS: Erchuan Zhang, David Suter, Ruwan Tennakoon, Tat-Jun Chin, Alireza Bab-Hadiashar, Giang Truong, Syed Zulqarnain Gilani
HIGHLIGHT: This paper studies the concept of weighted influences for solving MaxCon.

311, TITLE: TransforMatcher: Match-to-Match Attention for Semantic Correspondence
http://openaccess.thecvf.com//content/CVPR2022/html/Kim_TransforMatcher_Match-to-Match_Attention_for_Semantic_Correspondence_CVPR_2022_paper.html
AUTHORS: Seungwook Kim, Juhong Min, Minsu Cho
HIGHLIGHT: In this work, we introduce a strong semantic image matching learner, dubbed TransforMatcher, which builds on the success of transformer networks in vision domains.

312, TITLE: Robust Outlier Detection by De-Biasing VAE Likelihoods
http://openaccess.thecvf.com//content/CVPR2022/html/Chauhan_Robust_Outlier_Detection_by_De-Biasing_VAE_Likelihoods_CVPR_2022_paper.html
AUTHORS: Kushal Chauhan, Barath Mohan U, Pradeep Shenoy, Manish Gupta, Devarajan Sridharan
HIGHLIGHT: We propose novel analytical and algorithmic approaches to ameliorate key biases with VAE likelihoods.

313, TITLE: Contour-Hugging Heatmaps for Landmark Detection
http://openaccess.thecvf.com//content/CVPR2022/html/McCouat_Contour-Hugging_Heatmaps_for_Landmark_Detection_CVPR_2022_paper.html
AUTHORS: James McCouat, Irina Voiculescu
HIGHLIGHT: We propose an effective and easy-to-implement method for simultaneously performing landmark detection in images and obtaining an ingenious uncertainty measurement for each landmark.

314, TITLE: Voxel Field Fusion for 3D Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Voxel_Field_Fusion_for_3D_Object_Detection_CVPR_2022_paper.html
AUTHORS: Yanwei Li, Xiaojuan Qi, Yukang Chen, Liwei Wang, Zeming Li, Jian Sun, Jiaya Jia
HIGHLIGHT: In this work, we present a conceptually simple yet effective framework for cross-modality 3D object detection, named voxel field fusion.

315, TITLE: Divide and Conquer: Compositional Experts for Generalized Novel Class Discovery
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Divide_and_Conquer_Compositional_Experts_for_Generalized_Novel_Class_Discovery_CVPR_2022_paper.html
AUTHORS: Muli Yang, Yuehua Zhu, Jiaping Yu, Aming Wu, Cheng Deng
HIGHLIGHT: Despite showing competitive performance on novel classes, they fail to generalize to recognizing samples from both base and novel sets. In this paper, we focus on this generalized setting of NCD (GNCD), and propose to divide and conquer it with two groups of Compositional Experts (ComEx).

316, TITLE: Programmatic Concept Learning for Human Motion Description and Synthesis
http://openaccess.thecvf.com//content/CVPR2022/html/Kulal_Programmatic_Concept_Learning_for_Human_Motion_Description_and_Synthesis_CVPR_2022_paper.html
AUTHORS: Sumith Kulal, Jiayuan Mao, Alex Aiken, Jiajun Wu
HIGHLIGHT: We introduce Programmatic Motion Concepts, a hierarchical motion representation for human actions that captures both low level motion and high level description as motion concepts.

317, TITLE: Interpretable Part-Whole Hierarchies and Conceptual-Semantic Relationships in Neural Networks
http://openaccess.thecvf.com//content/CVPR2022/html/Garau_Interpretable_Part-Whole_Hierarchies_and_Conceptual-Semantic_Relationships_in_Neural_Networks_CVPR_2022_paper.html
AUTHORS: Nicola Garau, Niccolò Bisagno, Zeno Sambugaro, Nicola Conci
HIGHLIGHT: In this paper, we want to make a step forward towards interpretability in neural networks, providing new tools to interpret their behavior.

318, TITLE: Fast Algorithm for Low-Rank Tensor Completion in Delay-Embedded Space
http://openaccess.thecvf.com//content/CVPR2022/html/Yamamoto_Fast_Algorithm_for_Low-Rank_Tensor_Completion_in_Delay-Embedded_Space_CVPR_2022_paper.html
AUTHORS: Ryuki Yamamoto, Hidekata Hontani, Akira Imakura, Tatsuya Yokota

HIGHLIGHT: Recent studies have shown high completion performance with a relatively small window size, but experiments with large window sizes require huge amount of memory and cannot be easily calculated. In this study, we address this serious computational issue, and propose its fast and efficient algorithm.

319, TITLE: Panoptic, Instance and Semantic Relations: A Relational Context Encoder To Enhance Panoptic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Borse_Panoptic_Instance_and_Semantic_Relations_A_Relational_Context_Encoder_To_CVPR_2022_paper.html
AUTHORS: Shubhankar Borse, Hyojin Park, Hong Cai, Debasmit Das, Risheek Garrepalli, Fatih Porikli
HIGHLIGHT: This paper presents a novel framework to integrate both semantic and instance contexts for panoptic segmentation.

320, TITLE: Point2Seq: Detecting 3D Objects As Sequences
http://openaccess.thecvf.com//content/CVPR2022/html/Xue_Point2Seq_Detecting_3D_Objects_As_Sequences_CVPR_2022_paper.html
AUTHORS: Yujing Xue, Jiageng Mao, Minzhe Niu, Hang Xu, Michael Bi Mi, Wei Zhang, Xiaogang Wang, Xinchao Wang
HIGHLIGHT: We present a simple and effective framework, named Point2Seq, for 3D object detection from point clouds.

321, TITLE: Less Is More: Generating Grounded Navigation Instructions From Landmarks
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Less_Is_More_Generating_Grounded_Navigation_Instructions_From_Landmarks_CVPR_2022_paper.html
AUTHORS: Su Wang, Ceslee Montgomery, Jordi Orbay, Vighnesh Birodkar, Aleksandra Faust, Izzeddin Gur, Natasha Jaques, Austin Waters, Jason Baldridge, Peter Anderson
HIGHLIGHT: We study the automatic generation of navigation instructions from 360-degree images captured on indoor routes.

322, TITLE: Task-Adaptive Negative Envision for Few-Shot Open-Set Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Huang_Task-Adaptive_Negative_Envision_for_Few-Shot_Open-Set_Recognition_CVPR_2022_paper.html
AUTHORS: Shiyuan Huang, Jiawei Ma, Guangxing Han, Shih-Fu Chang
HIGHLIGHT: In this paper, we instead propose task-adaptive negative class envision for FSOR to integrate threshold tuning into the learning process.

323, TITLE: DisARM: Displacement Aware Relation Module for 3D Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Duan_DisARM_Displacement_Aware_Relation_Module_for_3D_Detection_CVPR_2022_paper.html
AUTHORS: Yao Duan, Chenyang Zhu, Yuqing Lan, Renjiao Yi, Xinwang Liu, Kai Xu
HIGHLIGHT: We introduce Displacement Aware Relation Module (DisARM), a novel neural network module for enhancing the performance of 3D object detection in point cloud scenes.

324, TITLE: ETHSeg: An Amodel Instance Segmentation Network and a Real-World Dataset for X-Ray Waste Inspection
http://openaccess.thecvf.com//content/CVPR2022/html/Qiu_ETHSeg_An_Amodel_Instance_Segmentation_Network_and_a_Real-World_Dataset_CVPR_2022_paper.html
AUTHORS: Lingteng Qiu, Zhangyang Xiong, Xuhao Wang, Kenkun Liu, Yihan Li, Guanying Chen, Xiaoguang Han, Shuguang Cui
HIGHLIGHT: We introduce a novel problem of instance-level waste segmentation in X-ray image for intelligent waste inspection, and contribute a real dataset consisting of 5,038 X-ray images (totally 30,881 waste items) with high-quality annotations (i.e., waste categories, object boxes, and instance-level masks) as a benchmark for this problem.

325, TITLE: MixFormer: Mixing Features Across Windows and Dimensions
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_MixFormer_Mixing_Features_Across_Windows_and_Dimensions_CVPR_2022_paper.html
AUTHORS: Qiang Chen, Qiman Wu, Jian Wang, Qinghao Hu, Tao Hu, Errui Ding, Jian Cheng, Jingdong Wang
HIGHLIGHT: While local-window self-attention performs notably in vision tasks, it suffers from limited receptive field and weak modeling capability issues. This is mainly because it performs self-attention within non-overlapped windows and shares weights on the channel dimension. We propose MixFormer to find a solution.

326, TITLE: Killing Two Birds With One Stone: Efficient and Robust Training of Face Recognition CNNs by Partial FC
http://openaccess.thecvf.com//content/CVPR2022/html/An_Killing_Two_Birds_With_One_Stone_Efficient_and_Robust_Training_CVPR_2022_paper.html
AUTHORS: Xiang An, Jiankang Deng, Jia Guo, Ziyong Feng, XuHan Zhu, Jing Yang, Tongliang Liu
HIGHLIGHT: In this paper, we propose a sparsely updating variant of the FC layer, named Partial FC (PFC).

327, TITLE: NeRF-Editing: Geometry Editing of Neural Radiance Fields
http://openaccess.thecvf.com//content/CVPR2022/html/Yuan_NeRF-Editing_Geometry_Editing_of_Neural_Radiance_Fields_CVPR_2022_paper.html
AUTHORS: Yu-Jie Yuan, Yang-Tian Sun, Yu-Kun Lai, Yuewen Ma, Rongfei Jia, Lin Gao
HIGHLIGHT: In this paper, we propose a method that allows users to perform controllable shape deformation on the implicit representation of the scene, and synthesizes the novel view images of the edited scene without re-training the network.

328, TITLE: Optimal Correction Cost for Object Detection Evaluation

http://openaccess.thecvf.com//content/CVPR2022/html/Otani_Optimal_Correction_Cost_for_Object_Detection_Evaluation_CVPR_2022_paper.html
AUTHORS: Mayu Otani, Riku Togashi, Yuta Nakashima, Esa Rahtu, Janne Heikkilä, Shin'ichi Satoh
HIGHLIGHT: To alleviate the gap between downstream tasks and the evaluation scenario, we propose Optimal Correction Cost (OC-cost), which assesses detection accuracy at image level.

329, TITLE: Contextual Similarity Distillation for Asymmetric Image Retrieval
http://openaccess.thecvf.com//content/CVPR2022/html/Wu_Contextual_Similarity_Distillation_for_Asymmetric_Image_Retrieval_CVPR_2022_paper.html
AUTHORS: Hui Wu, Min Wang, Wengang Zhou, Houqiang Li, Qi Tian
HIGHLIGHT: However, existing approaches either fail to achieve feature coherence or make strong assumptions, e.g., requiring labeled datasets or classifiers from large model, etc., which limits their practical application. To this end, we propose a flexible contextual similarity distillation framework to enhance the small query model and keep its output feature compatible with that of large gallery model, which is crucial with asymmetric retrieval.

330, TITLE: FineDiving: A Fine-Grained Dataset for Procedure-Aware Action Quality Assessment
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_FineDiving_A_Fine-Grained_Dataset_for_Procedure-Aware_Action_Quality_Assessment_CVPR_2022_paper.html
AUTHORS: Jinglin Xu, Yongming Rao, Xumin Yu, Guangyi Chen, Jie Zhou, Jiwen Lu
HIGHLIGHT: Specifically, we propose to parse pairwise query and exemplar action instances into consecutive steps with diverse semantic and temporal correspondences.

331, TITLE: Artistic Style Discovery With Independent Components
http://openaccess.thecvf.com//content/CVPR2022/html/Xie_Artistic_Style_Discovery_With_Independent_Components_CVPR_2022_paper.html
AUTHORS: Xin Xie, Yi Li, Huaibo Huang, Haiyan Fu, Wanwan Wang, Yanqing Guo
HIGHLIGHT: In particular, we take a closer look into the mechanism of style transfer and obtain different artistic style components from the latent space consisting of different style features.

332, TITLE: HEAT: Holistic Edge Attention Transformer for Structured Reconstruction
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_HEAT_Holistic_Edge_Attention_Transformer_for_Structured_Reconstruction_CVPR_2022_paper.html
AUTHORS: Jiacheng Chen, Yiming Qian, Yasutaka Furukawa
HIGHLIGHT: This paper presents a novel attention-based neural network for structured reconstruction, which takes a 2D raster image as an input and reconstructs a planar graph depicting an underlying geometric structure.

333, TITLE: HyperStyle: StyleGAN Inversion With HyperNetworks for Real Image Editing
http://openaccess.thecvf.com//content/CVPR2022/html/Alaluf_HyperStyle_StyleGAN_Inversion_With_HyperNetworks_for_Real_Image_Editing_CVPR_2022_paper.html
AUTHORS: Yuval Alaluf, Omer Tov, Ron Mokady, Rinon Gal, Amit Bermano
HIGHLIGHT: In this work, we introduce this approach into the realm of encoder-based inversion.

334, TITLE: DASO: Distribution-Aware Semantics-Oriented Pseudo-Label for Imbalanced Semi-Supervised Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Oh_DASO_Distribution-Aware_Semantics-Oriented_Pseudo-Label_for_Imbalanced_Semi-Supervised_Learning_CVPR_2022_paper.html
AUTHORS: Youngtaek Oh, Dong-Jin Kim, In So Kweon
HIGHLIGHT: The capability of the traditional semi-supervised learning (SSL) methods is far from real-world application due to severely biased pseudo-labels caused by (1) class imbalance and (2) class distribution mismatch between labeled and unlabeled data. This paper addresses such a relatively under-explored problem.

335, TITLE: Mobile-Former: Bridging MobileNet and Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Mobile-Former_Bridging_MobileNet_and_Transformer_CVPR_2022_paper.html
AUTHORS: Yinpeng Chen, Xiyang Dai, Dongdong Chen, Mengchen Liu, Xiaoyi Dong, Lu Yuan, Zicheng Liu
HIGHLIGHT: We present Mobile-Former, a parallel design of MobileNet and transformer with a two-way bridge in between.

336, TITLE: Exploiting Pseudo Labels in a Self-Supervised Learning Framework for Improved Monocular Depth Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Petrovai_Exploiting_Pseudo_Labels_in_a_Self-Supervised_Learning_Framework_for_Improved_CVPR_2022_paper.html
AUTHORS: Andra Petrovai, Sergiu Nedevschi
HIGHLIGHT: We present a novel self-distillation based self-supervised monocular depth estimation (SD-SSMDE) learning framework.

337, TITLE: DESTR: Object Detection With Split Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/He_DESTR_Object_Detection_With_Split_Transformer_CVPR_2022_paper.html
AUTHORS: Liqiang He, Sinisa Todorovic
HIGHLIGHT: First, we propose a new Detection Split Transformer (DESTR) that separates estimation of cross-attention into two independent branches -- one tailored for classification and the other for box regression.

338, TITLE: LTP: Lane-Based Trajectory Prediction for Autonomous Driving
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_LTP_Lane-
Based_Trajectory_Prediction_for_Autonomous_Driving_CVPR_2022_paper.html
AUTHORS: Jingke Wang, Tengju Ye, Ziqing Gu, Junbo Chen
HIGHLIGHT: This paper proposes a two-stage proposal-based motion forecasting method that exploits the sliced lane segments as fine-grained, shareable, and interpretable proposals.

339, TITLE: CycleMix: A Holistic Strategy for Medical Image Segmentation From Scribble Supervision
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_CycleMix_A_Holistic_Strategy_for_Medical_Image_Segmentation_From_Scribble_CVPR_2022_paper.html
AUTHORS: Ke Zhang, Xiahai Zhuang
HIGHLIGHT: To address the difficulties, we propose a new framework for scribble learning-based medical image segmentation, which is composed of mix augmentation and cycle consistency and thus is referred to as CycleMix.

340, TITLE: VideoINR: Learning Video Implicit Neural Representation for Continuous Space-Time Super-Resolution
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_VideoINR_Learning_Video_Implicit_Neural_Representation_for_Continuous_Space-Time_Super-Resolution_CVPR_2022_paper.html
AUTHORS: Zeyuan Chen, Yinbo Chen, Jingwen Liu, Xingqian Xu, Vidit Goel, Zhangyang Wang, Humphrey Shi, Xiaolong Wang
HIGHLIGHT: However, most of them only support a fixed up-sampling scale, which limits their flexibility and applications. In this work, instead of following the discrete representations, we propose Video Implicit Neural Representation (VideoINR), and we show its applications for STVSR.

341, TITLE: Towards End-to-End Unified Scene Text Detection and Layout Analysis
http://openaccess.thecvf.com//content/CVPR2022/html/Long_Towards_End-to-
End_Unified_Scene_Text_Detection_and_Layout_Analysis_CVPR_2022_paper.html
AUTHORS: Shangbang Long, Siyang Qin, Dmitry Panteleev, Alessandro Bissacco, Yasuhisa Fujii, Michalis Raptis
HIGHLIGHT: Scene text detection and document layout analysis have long been treated as two separate tasks in different image domains. In this paper, we bring them together and introduce the task of unified scene text detection and layout analysis.

342, TITLE: Image Based Reconstruction of Liquids From 2D Surface Detections
http://openaccess.thecvf.com//content/CVPR2022/html/Richter_Image_Based_Reconstruction_of_Liquids_From_2D_Surface_Detections_CVPR_2022_paper.html
AUTHORS: Florian Richter, Ryan K. Orosco, Michael C. Yip
HIGHLIGHT: In this work, we present a solution to the challenging problem of reconstructing liquids from image data.

343, TITLE: Contextual Outpainting With Object-Level Contrastive Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Contextual_Outpainting_With_Object-
Level_Contrastive_Learning_CVPR_2022_paper.html
AUTHORS: Jiacheng Li, Chang Chen, Zhiwei Xiong
HIGHLIGHT: To explore the semantic cues provided by the remaining foreground contents, we propose a novel ConTextual Outpainting GAN (CTO-GAN), leveraging the semantic layout as a bridge to synthesize coherent and diverse background contents.

344, TITLE: AP-BSN: Self-Supervised Denoising for Real-World Images via Asymmetric PD and Blind-Spot Network
http://openaccess.thecvf.com//content/CVPR2022/html/Lee_AP-BSN_Self-Supervised_Denoising_for_Real-
World_Images_via_Asymmetric_PD_and_CVPR_2022_paper.html
AUTHORS: Wooseok Lee, Sanghyun Son, Kyoung Mu Lee
HIGHLIGHT: However, it is not trivial to integrate PD and BSN directly, which prevents the fully self-supervised denoising model on real-world images. We propose an Asymmetric PD (AP) to address this issue, which introduces different PD stride factors for training and inference.

345, TITLE: AutoSDF: Shape Priors for 3D Completion, Reconstruction and Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Mittal_AutoSDF_Shape_Priors_for_3D_Completion_Reconstruction_and_Generation_CVPR_2022_paper.html
AUTHORS: Paritosh Mittal, Yen-Chi Cheng, Maneesh Singh, Shubham Tulsiani
HIGHLIGHT: In this paper, we propose an autoregressive prior for 3D shapes to solve multimodal 3D tasks such as shape completion, reconstruction, and generation.

346, TITLE: ISNAS-DIP: Image-Specific Neural Architecture Search for Deep Image Prior
http://openaccess.thecvf.com//content/CVPR2022/html/Arican_ISNAS-DIP_Image-
Specific_Neural_Architecture_Search_for_Deep_Image_Prior_CVPR_2022_paper.html
AUTHORS: Metin Ersin Arican, Ozgur Kara, Gustav Bredell, Ender Konukoglu
HIGHLIGHT: In this work, we first show that optimal neural architectures in the DIP framework are image-dependent. Leveraging this insight, we then propose an image-specific NAS strategy for the DIP framework that requires substantially less training than typical NAS approaches, effectively enabling image-specific NAS.

347, TITLE: Depth-Guided Sparse Structure-From-Motion for Movies and TV Shows
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Depth-Guided_Sparse_Structure-From-
Motion_for_Movies_and_TV_Shows_CVPR_2022_paper.html
AUTHORS: Sheng Liu, Xiaohan Nie, Raffay Hamid

HIGHLIGHT: The resulting small-motion parallax between video frames makes standard geometry-based SfM approaches not as effective for movies and TV shows. To address this challenge, we propose a simple yet effective approach that uses single-frame depth-prior obtained from a pretrained network to significantly improve geometry-based SfM for our small-parallax setting.

348, TITLE: End-to-End Referring Video Object Segmentation With Multimodal Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Botach_End-to-End_Referring_Video_Object_Segmentation_With_Multimodal_Transformers_CVPR_2022_paper.html
AUTHORS: Adam Botach, Evgenii Zheltonozhskii, Chaim Baskin
HIGHLIGHT: In this paper, we propose a simple Transformer-based approach to RVOS.

349, TITLE: Unpaired Cartoon Image Synthesis via Gated Cycle Mapping
http://openaccess.thecvf.com//content/CVPR2022/html/Men_Unpaired_Cartoon_Image_Synthesis_via_Gated_Cycle_Mapping_CVPR_2022_paper.html
AUTHORS: Yifang Men, Yuan Yao, Miaomiao Cui, Zhouhui Lian, Xuansong Xie, Xian-Sheng Hua
HIGHLIGHT: In this paper, we present a general-purpose solution to cartoon image synthesis with unpaired training data.

350, TITLE: IterMVS: Iterative Probability Estimation for Efficient Multi-View Stereo
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_IterMVS_Iterative_Probability_Estimation_for_Efficient_Multi-View_Stereo_CVPR_2022_paper.html
AUTHORS: Fangjinhua Wang, Silvano Galliani, Christoph Vogel, Marc Pollefeys
HIGHLIGHT: We present IterMVS, a new data-driven method for high-resolution multi-view stereo.

351, TITLE: Not All Points Are Equal: Learning Highly Efficient Point-Based Detectors for 3D LiDAR Point Clouds
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Not_All_Points_Are_Equal_Learning_Highly_Efficient_Point-Based_Detectors_CVPR_2022_paper.html
AUTHORS: Yifan Zhang, Qingyong Hu, Guoquan Xu, Yanxin Ma, Jianwei Wan, Yulan Guo
HIGHLIGHT: In particular, the foreground points are inherently more important than background points for object detectors. Motivated by this, we propose a highly-efficient single-stage point-based 3D detector in this paper, termed IA-SSD.

352, TITLE: FedCorr: Multi-Stage Federated Learning for Label Noise Correction
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_FedCorr_Multi-Stage_Federated_Learning_for_Label_Noise_Correction_CVPR_2022_paper.html
AUTHORS: Jingyi Xu, Zihan Chen, Tony Q.S. Quek, Kai Fong Ernest Chong
HIGHLIGHT: In this paper, we propose FedCorr, a general multi-stage framework to tackle heterogeneous label noise in FL, without making any assumptions on the noise models of local clients, while still maintaining client data privacy.

353, TITLE: Detecting Camouflaged Object in Frequency Domain
http://openaccess.thecvf.com//content/CVPR2022/html/Zhong_Detecting_Camouflaged_Object_in_Frequency_Domain_CVPR_2022_paper.html
AUTHORS: Yijie Zhong, Bo Li, Lv Tang, Senyun Kuang, Shuang Wu, Shouhong Ding
HIGHLIGHT: To well involve the frequency clues into the CNN models, we present a powerful network with two special components.

354, TITLE: RigNeRF: Fully Controllable Neural 3D Portraits
http://openaccess.thecvf.com//content/CVPR2022/html/Athar_RigNeRF_Fully_Controllable_Neural_3D_Portraits_CVPR_2022_paper.html
AUTHORS: ShahRukh Athar, Zexiang Xu, Kalyan Sunkavalli, Eli Shechtman, Zhixin Shu
HIGHLIGHT: In this work, we propose RigNeRF, a system that goes beyond just novel view synthesis and enables full control of head pose and facial expressions learned from a single portrait video.

355, TITLE: CLIP-Forge: Towards Zero-Shot Text-To-Shape Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Sanghi_CLIP-Forge_Towards_Zero-Shot_Text-To-Shape_Generation_CVPR_2022_paper.html
AUTHORS: Aditya Sanghi, Hang Chu, Joseph G. Lambourne, Ye Wang, Chin-Yi Cheng, Marco Fumero, Kamal Rahimi Malekshan
HIGHLIGHT: While significant recent progress has been made in text-to-image generation, text-to-shape generation remains a challenging problem due to the unavailability of paired text and shape data at a large scale. We present a simple yet effective method for zero-shot text-to-shape generation that circumvents such data scarcity.

356, TITLE: Style-Based Global Appearance Flow for Virtual Try-On
http://openaccess.thecvf.com//content/CVPR2022/html/He_Style-Based_Global_Appearance_Flow_for_Virtual_Try-On_CVPR_2022_paper.html
AUTHORS: Sen He, Yi-Zhe Song, Tao Xiang
HIGHLIGHT: They are thus intrinsically susceptible to difficult body poses/occlusions and large mis-alignments between person and garment images. To overcome this limitation, a novel global appearance flow estimation model is proposed in this work.

357, TITLE: Source-Free Object Detection by Learning To Overlook Domain Style
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Source-Free_Object_Detection_by_Learning_To_Overlook_Domain_Style_CVPR_2022_paper.html
AUTHORS: Shuaifeng Li, Mao Ye, Xiatian Zhu, Lihua Zhou, Lin Xiong

HIGHLIGHT: This approach suffers from both unsatisfactory accuracy of pseudo labels due to the presence of domain shift and limited use of target domain training data. In this work, we present a novel Learning to Overlook Domain Style (LODS) method with such limitations solved in a principled manner.

358, TITLE: Active Learning for Open-Set Annotation
http://openaccess.thecvf.com//content/CVPR2022/html/Ning_Active_Learning_for_Open-Set_Annotation_CVPR_2022_paper.html
AUTHORS: Kun-Peng Ning, Xun Zhao, Yu Li, Sheng-Jun Huang
HIGHLIGHT: However, in real annotation tasks, the unlabeled data usually contains a large amount of examples from unknown classes, resulting in the failure of most active learning methods. To tackle this open-set annotation (OSA) problem, we propose a new active learning framework called LfOSA, which boosts the classification performance with an effective sampling strategy to precisely detect examples from known classes for annotation.

359, TITLE: SceneSqueezer: Learning To Compress Scene for Camera Relocalization
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_SceneSqueezer_Learning_To_Compress_Scene_for_Camera_Relocalization_CVPR_2022_paper.html
AUTHORS: Luwei Yang, Rakesh Shrestha, Wenbo Li, Shuaicheng Liu, Guofeng Zhang, Zhaopeng Cui, Ping Tan
HIGHLIGHT: We design a novel framework that compresses a scene while still maintaining localization accuracy.

360, TITLE: SelfRecon: Self Reconstruction Your Digital Avatar From Monocular Video
http://openaccess.thecvf.com//content/CVPR2022/html/Jiang_SelfRecon_Self_Reconstruction_Your_Digital_Avatar_From_Monocular_Video_CVPR_2022_paper.html
AUTHORS: Boyi Jiang, Yang Hong, Hujun Bao, Juyong Zhang
HIGHLIGHT: We propose SelfRecon, a clothed human body reconstruction method that combines implicit and explicit representations to recover space-time coherent geometries from a monocular self-rotating human video.

361, TITLE: Instance-Dependent Label-Noise Learning With Manifold-Regularized Transition Matrix Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Cheng_Instance-Dependent_Label-Noise_Learning_With_Manifold-Regularized_Transition_Matrix_Estimation_CVPR_2022_paper.html
AUTHORS: De Cheng, Tongliang Liu, Yixiong Ning, Nannan Wang, Bo Han, Gang Niu, Xinbo Gao, Masashi Sugiyama
HIGHLIGHT: To address this problem, we have noticed that, there are psychological and physiological evidences showing that we humans are more likely to annotate instances of similar appearances to the same classes, and thus poor-quality or ambiguous instances of similar appearances are easier to be mislabeled to the correlated or same noisy classes. Therefore, we propose assumption on the geometry of $T(x)$ that "the closer two instances are, the more similar their corresponding transition matrices should be"..

362, TITLE: Rethinking the Augmentation Module in Contrastive Learning: Learning Hierarchical Augmentation Invariance With Expanded Views
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Rethinking_the_Augmentation_Module_in_Contrastive_Learning_Learning_Hierarchical_Augmentation_CVPR_2022_paper.html
AUTHORS: Junbo Zhang, Kaisheng Ma
HIGHLIGHT: Treating each type of augmentation equally during training makes the model learn non-optimal representations for various downstream tasks and limits the flexibility to choose augmentation types beforehand. Second, the strong data augmentations used in classic contrastive learning methods may bring too much invariance in some cases, and fine-grained information that is essential to some downstream tasks may be lost. This paper proposes a general method to alleviate these two problems by considering "where" and "what" to contrast in a general contrastive learning framework.

363, TITLE: Self-Supervised Models Are Continual Learners
http://openaccess.thecvf.com//content/CVPR2022/html/Fini_Self-Supervised_Models_Are_Continual_Learners_CVPR_2022_paper.html
AUTHORS: Enrico Fini, Victor G. Turrisi da Costa, Xavier Alameda-Pineda, Elisa Ricci, Karteek Alahari, Julien Mairal
HIGHLIGHT: In this paper, we show that self-supervised loss functions can be seamlessly converted into distillation mechanisms for CL by adding a predictor network that maps the current state of the representations to their past state.

364, TITLE: Dreaming To Prune Image Deraining Networks
http://openaccess.thecvf.com//content/CVPR2022/html/Zou_Dreaming_To_Prune_Image_Deraining_Networks_CVPR_2022_paper.html
AUTHORS: Weiqi Zou, Yang Wang, Xueyang Fu, Yang Cao
HIGHLIGHT: We note that it is overstretched to fine-tune the compressed model using self-collected data, as it exhibits poor generalization over images with different degradation characteristics. To address this problem, we propose a novel data-free compression framework for deraining networks.

365, TITLE: Equivariant Point Cloud Analysis via Learning Orientations for Message Passing
http://openaccess.thecvf.com//content/CVPR2022/html/Luo_Equivariant_Point_Cloud_Analysis_via_Learning_Orientations_for_Message_Passing_CVPR_2022_paper.html
AUTHORS: Shitong Luo, Jiahan Li, Jiaqi Guan, Yufeng Su, Chaoran Cheng, Jian Peng, Jianzhu Ma
HIGHLIGHT: In this work, we propose a novel and simple framework to achieve equivariance for point cloud analysis based on the message passing (graph neural network) scheme.

366, TITLE: When Does Contrastive Visual Representation Learning Work?
http://openaccess.thecvf.com//content/CVPR2022/html/Cole_When_Does_Contrastive_Visual_Representation_Learning_Work_CVPR_2022_paper.html

AUTHORS: Elijah Cole, Xuan Yang, Kimberly Wilber, Oisin Mac Aodha, Serge Belongie
HIGHLIGHT: By looking through the lenses of data quantity, data domain, data quality, and task granularity, we provide new insights into the necessary conditions for successful self-supervised learning.

367, TITLE: One Step at a Time: Long-Horizon Vision-and-Language Navigation With Milestones
http://openaccess.thecvf.com//content/CVPR2022/html/Song_One_Step_at_a_Time_Long-Horizon_Vision-and-Language_Navigation_With_Milestones_CVPR_2022_paper.html
AUTHORS: Chan Hee Song, Jihyung Kil, Tai-Yu Pan, Brian M. Sadler, Wei-Lun Chao, Yu Su
HIGHLIGHT: However, when it comes to long-horizon tasks with extended sequences of actions, an agent can easily ignore some instructions or get stuck in the middle of the long instructions and eventually fail the task. To address this challenge, we propose a model-agnostic milestone-based task tracker (M-Track) to guide the agent and monitor its progress.

368, TITLE: Node Representation Learning in Graph via Node-to-Neighbourhood Mutual Information Maximization
http://openaccess.thecvf.com//content/CVPR2022/html/Dong_Node_Representation_Learning_in_Graph_via_Node-to-Neighbourhood_Mutual_Information_Maximization_CVPR_2022_paper.html
AUTHORS: Wei Dong, Junsheng Wu, Yi Luo, Zongyuan Ge, Peng Wang
HIGHLIGHT: In this work, we present a simple-yet-effective self-supervised node representation learning strategy via directly maximizing the mutual information between the hidden representations of nodes and their neighbourhood, which can be theoretically justified by its link to graph smoothing.

369, TITLE: Point Cloud Pre-Training With Natural 3D Structures
http://openaccess.thecvf.com//content/CVPR2022/html/Yamada_Point_Cloud_Pre-Training_With_Natural_3D_Structures_CVPR_2022_paper.html
AUTHORS: Ryosuke Yamada, Hirokatsu Kataoka, Naoya Chiba, Yukiyasu Domae, Tetsuya Ogata
HIGHLIGHT: Therefore, constructing a largescale 3D point clouds dataset is difficult. In order to remedy this issue, we propose a newly developed point cloud fractal database (PC-FractalDB), which is a novel family of formula-driven supervised learning inspired by fractal geometry encountered in natural 3D structures.

370, TITLE: Scene Consistency Representation Learning for Video Scene Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Wu_Scene_Consistency_Representation_Learning_for_Video_Scene_Segmentation_CVPR_2022_paper.html
AUTHORS: Haoqian Wu, Keyu Chen, Yanan Luo, Ruizhi Qiao, Bo Ren, Haozhe Liu, Weicheng Xie, Linlin Shen
HIGHLIGHT: Spotting the correct scene boundary from the long-term video is a challenging task, since a model must understand the storyline of the video to figure out where a scene starts and ends. To this end, we propose an effective Self-Supervised Learning (SSL) framework to learn better shot representations from unlabeled long-term videos.

371, TITLE: Two Coupled Rejection Metrics Can Tell Adversarial Examples Apart
http://openaccess.thecvf.com//content/CVPR2022/html/Pang_Two_Coupled_Rejection_Metrics_Can_Tell_Adversarial_Examples_Apart_CVPR_2022_paper.html
AUTHORS: Tianyu Pang, Huishuai Zhang, Di He, Yinpeng Dong, Hang Su, Wei Chen, Jun Zhu, Tie-Yan Liu
HIGHLIGHT: A complementary way towards robustness is to introduce a rejection option, allowing the model to not return predictions on uncertain inputs, where confidence is a commonly used certainty proxy. Along with this routine, we find that confidence and a rectified confidence (R-Con) can form two coupled rejection metrics, which could provably distinguish wrongly classified inputs from correctly classified ones.

372, TITLE: Exploiting Explainable Metrics for Augmented SGD
http://openaccess.thecvf.com//content/CVPR2022/html/Hosseini_Exploiting_Explainable_Metrics_for_Augmented_SGD_CVPR_2022_paper.html
AUTHORS: Mahdi S. Hosseini, Mathieu Tuli, Konstantinos N. Plataniotis
HIGHLIGHT: In this paper, we address the following question: can we probe intermediate layers of a deep neural network to identify and quantify the learning quality of each layer?

373, TITLE: Semi-Supervised Video Semantic Segmentation With Inter-Frame Feature Reconstruction
http://openaccess.thecvf.com//content/CVPR2022/html/Zhuang_Semi-Supervised_Video_Semantic_Segmentation_With_Inter-Frame_Feature_Reconstruction_CVPR_2022_paper.html
AUTHORS: Jiafan Zhuang, Zilei Wang, Yuan Gao
HIGHLIGHT: This is called inner-video overfitting, and it would actually lead to inferior performance. To tackle this issue, we propose a novel inter-frame feature reconstruction (IFR) technique to leverage the ground-truth labels to supervise the model training on unlabeled frames.

374, TITLE: GenDR: A Generalized Differentiable Renderer
http://openaccess.thecvf.com//content/CVPR2022/html/Petersen_GenDR_A_Generalized_Differentiable_Renderer_CVPR_2022_paper.html
AUTHORS: Felix Petersen, Bastian Goldluecke, Christian Borgelt, Oliver Deussen
HIGHLIGHT: In this work, we present and study a generalized family of differentiable renderers.

375, TITLE: Improving Neural Implicit Surfaces Geometry With Patch Warping
http://openaccess.thecvf.com//content/CVPR2022/html/Darmon_Improving_Neural_Implicit_Surfaces_Geometry_With_Patch_Warping_CVPR_2022_paper.html

AUTHORS:		François Darmon, Bénédicte Bascle, Jean-Clément Devaux, Pascal Monasse, Mathieu Aubry
HIGHLIGHT:		Neural implicit surfaces have become an important technique for multi-view 3D reconstruction but their accuracy remains limited. In this paper, we argue that this comes from the difficulty to learn and render high frequency textures with neural networks.

376, TITLE:		XYLayoutLM: Towards Layout-Aware Multimodal Networks for Visually-Rich Document Understanding
http://openaccess.thecvf.com//content/CVPR2022/html/Gu_XYLayoutLM_Towards_Layout-Aware_Multimodal_Networks_for_Visually-Rich_Document_Understanding_CVPR_2022_paper.html
AUTHORS:		Zhangxuan Gu, Changhua Meng, Ke Wang, Jun Lan, Weiqiang Wang, Ming Gu, Liqing Zhang
HIGHLIGHT:		In this paper, we propose a robust layout-aware multimodal network named XYLayoutLM to capture and leverage rich layout information from proper reading orders produced by our Augmented XY Cut.

377, TITLE:		Amodal Segmentation Through Out-of-Task and Out-of-Distribution Generalization With a Bayesian Model
http://openaccess.thecvf.com//content/CVPR2022/html/Sun_Amodal_Segmentation_Through_Out-of-Task_and_Out-of-Distribution_Generalization_With_a_Bayesian_CVPR_2022_paper.html
AUTHORS:		Yihong Sun, Adam Kortylewski, Alan Yuille
HIGHLIGHT:		This task is particularly challenging for deep neural networks because data is difficult to obtain and annotate. Therefore, we formulate amodal segmentation as an out-of-task and out-of-distribution generalization problem.

378, TITLE:		How Well Do Sparse ImageNet Models Transfer?
http://openaccess.thecvf.com//content/CVPR2022/html/Iofinova_How_Well_Do_Sparse_ImageNet_Models_Transfer_CVPR_2022_paper.html
AUTHORS:		Eugenia Iofinova, Alexandra Peste, Mark Kurtz, Dan Alistarh
HIGHLIGHT:		Generally, more accurate models on the "upstream" dataset tend to provide better transfer accuracy "downstream". In this work, we perform an in-depth investigation of this phenomenon in the context of convolutional neural networks (CNNs) trained on the ImageNet dataset, which have been pruned--that is, compressed by sparsifiying their connections.

379, TITLE:		REX: Reasoning-Aware and Grounded Explanation
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_REX_Reasoning-Aware_and_Grounded_Explanation_CVPR_2022_paper.html
AUTHORS:		Shi Chen, Qi Zhao
HIGHLIGHT:		This paper aims to close the gap from three distinct perspectives: first, we define a new type of multi-modal explanations that explain the decisions by progressively traversing the reasoning process and grounding keywords in the images.

380, TITLE:		Dynamic Dual-Output Diffusion Models
http://openaccess.thecvf.com//content/CVPR2022/html/Benny_Dynamic_Dual-Output_Diffusion_Models_CVPR_2022_paper.html
AUTHORS:		Yaniv Benny, Lior Wolf
HIGHLIGHT:		In this paper, we reveal some of the causes that affect the generation quality of diffusion models, especially when sampling with few iterations, and come up with a simple, yet effective, solution to mitigate them.

381, TITLE:		StyleT2I: Toward Compositional and High-Fidelity Text-to-Image Synthesis
http://openaccess.thecvf.com//content/CVPR2022/html/Li_StyleT2I_Toward_Compositional_and_High-Fidelity_Text-to-Image_Synthesis_CVPR_2022_paper.html
AUTHORS:		Zhiheng Li, Martin Renqiang Min, Kai Li, Chenliang Xu
HIGHLIGHT:		In this paper, we introduce a new framework, StyleT2I, to improve the compositionality of text-to-image synthesis.

382, TITLE:		JoinABLe: Learning Bottom-Up Assembly of Parametric CAD Joints
http://openaccess.thecvf.com//content/CVPR2022/html/Willis_JoinABLe_Learning_Bottom-Up_Assembly_of_Parametric_CAD_Joints_CVPR_2022_paper.html
AUTHORS:		Karl D.D. Willis, Pradeep Kumar Jayaraman, Hang Chu, Yunsheng Tian, Yifei Li, Daniele Grandi, Aditya Sanghi, Linh Tran, Joseph G. Lambourne, Armando Solar-Lezama, Wojciech Matusik
HIGHLIGHT:		In this paper we introduce JoinABLe, a learning-based method that assembles parts together to form joints.

383, TITLE:		CaDeX: Learning Canonical Deformation Coordinate Space for Dynamic Surface Representation via Neural Homeomorphism
http://openaccess.thecvf.com//content/CVPR2022/html/Lei_CaDeX_Learning_Canonical_Deformation_Coordinate_Space_for_Dynamic_Surface_Representation_CVPR_2022_paper.html
AUTHORS:		Jiahui Lei, Kostas Daniilidis
HIGHLIGHT:		We introduce Canonical Deformation Coordinate Space (CaDeX), a unified representation of both shape and nonrigid motion.

384, TITLE:		Canonical Voting: Towards Robust Oriented Bounding Box Detection in 3D Scenes
http://openaccess.thecvf.com//content/CVPR2022/html/You_Canonical_Voting_Towards_Robust_Oriented_Bounding_Box_Detection_in_3D_CVPR_2022_paper.html
AUTHORS:		Yang You, Zelin Ye, Yujing Lou, Chengkun Li, Yong-Lu Li, Lizhuang Ma, Weiming Wang, Cewu Lu
HIGHLIGHT:		In the work, we disentangle the direct offset into Local Canonical Coordinates (LCC), box scales and box orientations.

385, TITLE: V-Doc: Visual Questions Answers With Documents
http://openaccess.thecvf.com//content/CVPR2022/html/Ding_V-
Doc_Visual_Questions_Answers_With_Documents_CVPR_2022_paper.html
AUTHORS: Yihao Ding, Zhe Huang, Runlin Wang, YanHang Zhang, Xianru Chen, Yuzhong Ma, Hyunsuk Chung, Soyeon Caren Han
HIGHLIGHT: We propose V-Doc, a question-answering tool using document images and PDF, mainly for researchers and general non-deep learning experts looking to generate, process, and understand the document visual question answering tasks.

386, TITLE: AEGNN: Asynchronous Event-Based Graph Neural Networks
http://openaccess.thecvf.com//content/CVPR2022/html/Schaefer_AEGNN_Asynchronous_Event-
Based_Graph_Neural_Networks_CVPR_2022_paper.html
AUTHORS: Simon Schaefer, Daniel Gehrig, Davide Scaramuzza
HIGHLIGHT: For this reason, recent works have adopted Graph Neural Networks (GNNs), which process events as "static" spatio-temporal graphs, which are inherently "sparse". We take this trend one step further by introducing Asynchronous, Event-based Graph Neural Networks (AEGNNs), a novel event-processing paradigm that generalizes standard GNNs to process events as "evolving" spatio-temporal graphs.

387, TITLE: Layer-Wised Model Aggregation for Personalized Federated Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Ma_Layer-
Wised_Model_Aggregation_for_Personalized_Federated_Learning_CVPR_2022_paper.html
AUTHORS: Xiaosong Ma, Jie Zhang, Song Guo, Wenchao Xu
HIGHLIGHT: In this paper, we propose a novel pFL training framework dubbed Layer-wised Personalized Federated learning (pFedLA) that can discern the importance of each layer from different clients, and thus is able to optimize the personalized model aggregation for clients with heterogeneous data.

388, TITLE: Polarity Sampling: Quality and Diversity Control of Pre-Trained Generative Networks via Singular Values
http://openaccess.thecvf.com//content/CVPR2022/html/Humayun_Polarity_Sampling_Quality_and_Diversity_Control_of_Pre-
Trained_Generative_Networks_CVPR_2022_paper.html
AUTHORS: Ahmed Imtiaz Humayun, Randall Balestriero, Richard Baraniuk
HIGHLIGHT: We present Polarity Sampling, a theoretically justified plug-and-play method for controlling the generation quality and diversity of any pre-trained deep generative network (DGN).

389, TITLE: Style-Structure Disentangled Features and Normalizing Flows for Diverse Icon Colorization
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Style-
Structure_Disentangled_Features_and_Normalizing_Flows_for_Diverse_Icon_Colorization_CVPR_2022_paper.html
AUTHORS: Yuan-kui Li, Yun-Hsuan Lien, Yu-Shuen Wang
HIGHLIGHT: In this study, we present a colorization network that generates flat-color icons according to given sketches and semantic colorization styles.

390, TITLE: Object-Aware Video-Language Pre-Training for Retrieval
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Object-Aware_Video-Language_Pre-
Training_for_Retrieval_CVPR_2022_paper.html
AUTHORS: Jinpeng Wang, Yixiao Ge, Guanyu Cai, Rui Yan, Xudong Lin, Ying Shan, Xiaohu Qie, Mike Zheng Shou
HIGHLIGHT: In this work, we present Object-aware Transformers, an object-centric approach that extends video-language transformer to incorporate object representations.

391, TITLE: OSKDet: Orientation-Sensitive Keypoint Localization for Rotated Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Lu_OSKDet_Orientation-
Sensitive_Keypoint_Localization_for_Rotated_Object_Detection_CVPR_2022_paper.html
AUTHORS: Dongchen Lu, Dongmei Li, Yali Li, Shengjin Wang
HIGHLIGHT: In this paper, we propose an orientation-sensitive keypoint based rotated detector OSKDet.

392, TITLE: MAT: Mask-Aware Transformer for Large Hole Image Inpainting
http://openaccess.thecvf.com//content/CVPR2022/html/Li_MAT_Mask-
Aware_Transformer_for_Large_Hole_Image_Inpainting_CVPR_2022_paper.html
AUTHORS: Wenbo Li, Zhe Lin, Kun Zhou, Lu Qi, Yi Wang, Jiaya Jia
HIGHLIGHT: In this paper, we present a novel transformer-based model for large hole inpainting, which unifies the merits of transformers and convolutions to efficiently process high-resolution images.

393, TITLE: Exploring Geometric Consistency for Monocular 3D Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Lian_Exploring_Geometric_Consistency_for_Monocular_3D_Object_Detecti
on_CVPR_2022_paper.html
AUTHORS: Qing Lian, Botao Ye, Ruijia Xu, Weilong Yao, Tong Zhang
HIGHLIGHT: In particular, we design a series of geometric manipulations to diagnose existing detectors and then illustrate their vulnerability to consistently associate the depth with object apparent sizes and positions. To alleviate this issue, we propose four geometry-aware data augmentation approaches to enhance the geometric consistency of the detectors.

394, TITLE: Neural Window Fully-Connected CRFs for Monocular Depth Estimation

http://openaccess.thecvf.com//content/CVPR2022/html/Yuan_Neural_Window_Fully-
Connected_CRFs_for_Monocular_Depth_Estimation_CVPR_2022_paper.html
AUTHORS: Weihao Yuan, Xiaodong Gu, Zuozhuo Dai, Siyu Zhu, Ping Tan
HIGHLIGHT: While recent works design increasingly complicated and powerful networks to directly regress the depth map,
we take the path of CRFs optimization.

395, TITLE: CodedVTR: Codebook-Based Sparse Voxel Transformer With Geometric Guidance
http://openaccess.thecvf.com//content/CVPR2022/html/Zhao_CodedVTR_Codebook-
Based_Sparse_Voxel_Transformer_With_Geometric_Guidance_CVPR_2022_paper.html
AUTHORS: Tianchen Zhao, Niansong Zhang, Xuefei Ning, He Wang, Li Yi, Yu Wang
HIGHLIGHT: We propose Codebook-based Voxel TRansformer), which improves data efficiency and generalization ability
for 3D sparse voxel transformers.

396, TITLE: Uncertainty-Aware Deep Multi-View Photometric Stereo
http://openaccess.thecvf.com//content/CVPR2022/html/Kaya_Uncertainty-Aware_Deep_Multi-
View_Photometric_Stereo_CVPR_2022_paper.html
AUTHORS: Berk Kaya, Suryansh Kumar, Carlos Oliveira, Vittorio Ferrari, Luc Van Gool
HIGHLIGHT: This paper presents a simple and effective solution to the longstanding classical multi-view photometric stereo
(MVPS) problem.

397, TITLE: Coherent Point Drift Revisited for Non-Rigid Shape Matching and Registration
http://openaccess.thecvf.com//content/CVPR2022/html/Fan_Coherent_Point_Drift_Revisited_for_Non-
Rigid_Shape_Matching_and_Registration_CVPR_2022_paper.html
AUTHORS: Aoxiang Fan, Jiayi Ma, Xin Tian, Xiaoguang Mei, Wei Liu
HIGHLIGHT: In this paper, we explore a new type of extrinsic method to directly align two geometric shapes with point-to-
point correspondences in ambient space by recovering a deformation, which allows more continuous and smooth maps to be obtained.

398, TITLE: Unleashing Potential of Unsupervised Pre-Training With Intra-Identity Regularization for Person Re-
Identification
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Unleashing_Potential_of_Unsupervised_Pre-Training_With_Intra-
Identity_Regularization_for_Person_CVPR_2022_paper.html
AUTHORS: Zizheng Yang, Xin Jin, Kecheng Zheng, Feng Zhao
HIGHLIGHT: Inspired by the great success of self-supervised representation learning with contrastive objectives, in this paper,
we design an Unsupervised Pre-training framework for ReID based on the contrastive learning (CL) pipeline, dubbed UP-ReID.

399, TITLE: Align and Prompt: Video-and-Language Pre-Training With Entity Prompts
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Align_and_Prompt_Video-and-Language_Pre-
Training_With_Entity_Prompts_CVPR_2022_paper.html
AUTHORS: Dongxu Li, Junnan Li, Hongdong Li, Juan Carlos Niebles, Steven C.H. Hoi
HIGHLIGHT: We propose Align and Prompt: an efficient and effective video-and-language pre-training framework with
better cross-modal alignment.

400, TITLE: A Unified Query-Based Paradigm for Point Cloud Understanding
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_A_Unified_Query-
Based_Paradigm_for_Point_Cloud_Understanding_CVPR_2022_paper.html
AUTHORS: Zetong Yang, Li Jiang, Yanan Sun, Bernt Schiele, Jiaya Jia
HIGHLIGHT: In this paper, we present a novel Embedding-Querying paradigm (EQ- Paradigm) for 3D understanding tasks
including detection, segmentation and classification.

401, TITLE: It's About Time: Analog Clock Reading in the Wild
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Its_About_Time_Analog_Clock_Reading_in_the_Wild_CVPR_2022_p
aper.html
AUTHORS: Charig Yang, Weidi Xie, Andrew Zisserman
HIGHLIGHT: In this paper, we present a framework for reading analog clocks in natural images or videos.

402, TITLE: MSG-Transformer: Exchanging Local Spatial Information by Manipulating Messenger Tokens
http://openaccess.thecvf.com//content/CVPR2022/html/Fang_MSG-
Transformer_Exchanging_Local_Spatial_Information_by_Manipulating_Messenger_Tokens_CVPR_2022_paper.html
AUTHORS: Jiemin Fang, Lingxi Xie, Xinggang Wang, Xiaopeng Zhang, Wenyu Liu, Qi Tian
HIGHLIGHT: This paper aims to alleviate the conflict between efficiency and flexibility, for which we propose a specialized
token for each region that serves as a messenger (MSG).

403, TITLE: Cross Modal Retrieval With Querybank Normalisation
http://openaccess.thecvf.com//content/CVPR2022/html/Bogolin_Cross_Modal_Retrieval_With_Querybank_Normalisation_CVPR_20
22_paper.html
AUTHORS: Simion-Vlad Bogolin, Ioana Croitoru, Hailin Jin, Yang Liu, Samuel Albanie
HIGHLIGHT: Drawing inspiration from the NLP literature, we formulate a simple but effective framework called Querybank
Normalisation (QB-Norm) that re-normalises query similarities to account for hubs in the embedding space.

404, TITLE: Contrastive Dual Gating: Learning Sparse Features With Contrastive Learning

http://openaccess.thecvf.com//content/CVPR2022/html/Meng_Contrastive_Dual_Gating_Learning_Sparse_Features_With_Contrastive_Learning_CVPR_2022_paper.html
AUTHORS: Jian Meng, Li Yang, Jinwoo Shin, Deliang Fan, Jae-sun Seo
HIGHLIGHT: However, we found that such salience predictors cannot be easily trained when they are naively applied to contrastive learning from scratch. To address this issue, we propose contrastive dual gating(CDG), a novel dynamic pruning algorithm that skips the uninformative features during contrastive learning without hurting the trainability of the networks.

405, TITLE: Universal Photometric Stereo Network Using Global Lighting Contexts
http://openaccess.thecvf.com//content/CVPR2022/html/Ikehata_Universal_Photometric_Stereo_Network_Using_Global_Lighting_Contexts_CVPR_2022_paper.html
AUTHORS: Satoshi Ikehata
HIGHLIGHT: Unlike existing tasks that assumed specific physical lighting models; hence, drastically limited their usability, a solution algorithm of this task is supposed to work for objects with diverse shapes and materials under arbitrary lighting variations without assuming any specific models. To solve this extremely challenging task, we present a purely data-driven method, which eliminates the prior assumption of lighting by replacing the recovery of physical lighting parameters with the extraction of the generic lighting representation, named global lighting contexts.

406, TITLE: Hire-MLP: Vision MLP via Hierarchical Rearrangement
http://openaccess.thecvf.com//content/CVPR2022/html/Guo_Hire-MLP_Vision_MLP_via_Hierarchical_Rearrangement_CVPR_2022_paper.html
AUTHORS: Jianyuan Guo, Yehui Tang, Kai Han, Xinghao Chen, Han Wu, Chao Xu, Chang Xu, Yunhe Wang
HIGHLIGHT: This paper presents Hire-MLP, a simple yet competitive vision MLP architecture via Hierarchical rearrangement, which contains two levels of rearrangements.

407, TITLE: Ray3D: Ray-Based 3D Human Pose Estimation for Monocular Absolute 3D Localization
http://openaccess.thecvf.com//content/CVPR2022/html/Zhan_Ray3D_Ray-Based_3D_Human_Pose_Estimation_for_Monocular_Absolute_3D_CVPR_2022_paper.html
AUTHORS: Yu Zhan, Fenghai Li, Renliang Weng, Wongun Choi
HIGHLIGHT: In this paper, we propose a novel monocular ray-based 3D (Ray3D) absolute human pose estimation with calibrated camera.

408, TITLE: Occluded Human Mesh Recovery
http://openaccess.thecvf.com//content/CVPR2022/html/Khirodkar_Occluded_Human_Mesh_Recovery_CVPR_2022_paper.html
AUTHORS: Rawal Khirodkar, Shashank Tripathi, Kris Kitani
HIGHLIGHT: Consequently, top-down methods have difficulties in recovering accurate 3D human meshes under severe person-person occlusion. To address this, we present Occluded Human Mesh Recovery (OCHMR) - a novel top-down mesh recovery approach that incorporates image spatial context to overcome the limitations of the single-human assumption.

409, TITLE: Multi-Object Tracking Meets Moving UAV
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Multi-Object_Tracking_Meets_Moving_UAV_CVPR_2022_paper.html
AUTHORS: Shuai Liu, Xin Li, Huchuan Lu, You He
HIGHLIGHT: In this paper, we propose a UAVMOT network specially for multi-object tracking in UAV views.

410, TITLE: ASM-Loc: Action-Aware Segment Modeling for Weakly-Supervised Temporal Action Localization
http://openaccess.thecvf.com//content/CVPR2022/html/He_ASM-Loc_Action-Aware_Segment_Modeling_for_Weakly-Supervised_Temporal_Action_Localization_CVPR_2022_paper.html
AUTHORS: Bo He, Xitong Yang, Le Kang, Zhiyu Cheng, Xin Zhou, Abhinav Shrivastava
HIGHLIGHT: However, this formulation typically treats snippets in a video as independent instances, ignoring the underlying temporal structures within and across action segments. To address this problem, we propose \system, a novel WTAL framework that enables explicit, action-aware segment modeling beyond standard MIL-based methods.

411, TITLE: Uncertainty-Guided Probabilistic Transformer for Complex Action Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Guo_Uncertainty-Guided_Probabilistic_Transformer_for_Complex_Action_Recognition_CVPR_2022_paper.html
AUTHORS: Hongji Guo, Hanjing Wang, Qiang Ji
HIGHLIGHT: This paper introduces a probabilistic model named Uncertainty-Guided Probabilistic Transformer (UGPT) for complex action recognition.

412, TITLE: Scaling Up Your Kernels to 31x31: Revisiting Large Kernel Design in CNNs
http://openaccess.thecvf.com//content/CVPR2022/html/Ding_Scaling_Up_Your_Kernels_to_31x31_Revisiting_Large_Kernel_Design_CVPR_2022_paper.html
AUTHORS: Xiaohan Ding, Xiangyu Zhang, Jungong Han, Guiguang Ding
HIGHLIGHT: Following the guidelines, we propose RepLKNet, a pure CNN architecture whose kernel size is as large as 31x31, in contrast to commonly used 3x3.

413, TITLE: End-to-End Multi-Person Pose Estimation With Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Shi_End-to-End_Multi-Person_Pose_Estimation_With_Transformers_CVPR_2022_paper.html
AUTHORS: Dahu Shi, Xing Wei, Liangqi Li, Ye Ren, Wenming Tan

HIGHLIGHT: In this paper, we propose the first fully end-to-end multi-person Pose Estimation framework with TRansformers, termed PETR.

414, TITLE: REGTR: End-to-End Point Cloud Correspondences With Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Yew_REGTR_End-to-End_Point_Cloud_Correspondences_With_Transformers_CVPR_2022_paper.html
AUTHORS: Zi Jian Yew, Gim Hee Lee
HIGHLIGHT: In this work, we conjecture that attention mechanisms can replace the role of explicit feature matching and RANSAC, and thus propose an end-to-end framework to directly predict the final set of correspondences.

415, TITLE: Neural 3D Scene Reconstruction With the Manhattan-World Assumption
http://openaccess.thecvf.com//content/CVPR2022/html/Guo_Neural_3D_Scene_Reconstruction_With_the_Manhattan-World_Assumption_CVPR_2022_paper.html
AUTHORS: Haoyu Guo, Sida Peng, Haotong Lin, Qianqian Wang, Guofeng Zhang, Hujun Bao, Xiaowei Zhou
HIGHLIGHT: In this work, we show that the planar constraints can be conveniently integrated into the recent implicit neural representation-based reconstruction methods.

416, TITLE: V2C: Visual Voice Cloning
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_V2C_Visual_Voice_Cloning_CVPR_2022_paper.html
AUTHORS: Qi Chen, Mingkui Tan, Yuankai Qi, Jiaqiu Zhou, Yuanqing Li, Qi Wu
HIGHLIGHT: However, there also exist many scenarios that cannot be well reflected by these VC tasks, such as movie dubbing, which requires the speech to be with emotions consistent with the movie plots. To fill this gap, in this work we propose a new task named Visual Voice Cloning (V2C), which seeks to convert a paragraph of text to a speech with both desired voice specified by a reference audio and desired emotion specified by a reference video.

417, TITLE: Revisiting AP Loss for Dense Object Detection: Adaptive Ranking Pair Selection
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_Revisiting_AP_Loss_for_Dense_Object_Detection_Adaptive_Ranking_Pair_CVPR_2022_paper.html
AUTHORS: Dongli Xu, Jinhong Deng, Wen Li
HIGHLIGHT: In this work, we revisit the average precision (AP) loss and reveal that the crucial element is that of selecting the ranking pairs between positive and negative samples.

418, TITLE: 3DeformRS: Certifying Spatial Deformations on Point Clouds
http://openaccess.thecvf.com//content/CVPR2022/html/S._3DeformRS_Certifying_Spatial_Deformations_on_Point_Clouds_CVPR_2022_paper.html
AUTHORS: Gabriel Pérez S., Juan C. Pérez, Motasem Alfarra, Silvio Giancola, Bernard Ghanem
HIGHLIGHT: In this work, we propose 3DeformRS, a method to certify the robustness of point cloud Deep Neural Networks (DNNs) against real-world deformations.

419, TITLE: ElePose: Unsupervised 3D Human Pose Estimation by Predicting Camera Elevation and Learning Normalizing Flows on 2D Poses
http://openaccess.thecvf.com//content/CVPR2022/html/Wandt_ElePose_Unsupervised_3D_Human_Pose_Estimation_by_Predicting_Camera_Elevation_CVPR_2022_paper.html
AUTHORS: Bastian Wandt, James J. Little, Helge Rhodin
HIGHLIGHT: Unfortunately, labeled training data does not yet exist for many human activities since 3D annotation requires dedicated motion capture systems. Therefore, we propose an unsupervised approach that learns to predict a 3D human pose from a single image while only being trained with 2D pose data, which can be crowd-sourced and is already widely available.

420, TITLE: MAD: A Scalable Dataset for Language Grounding in Videos From Movie Audio Descriptions
http://openaccess.thecvf.com//content/CVPR2022/html/Soldan_MAD_A_Scalable_Dataset_for_Language_Grounding_in_Videos_From_CVPR_2022_paper.html
AUTHORS: Mattia Soldan, Alejandro Pardo, Juan León Alcázar, Fabian Caba, Chen Zhao, Silvio Giancola, Bernard Ghanem
HIGHLIGHT: In this work, we present MAD (Movie Audio Descriptions), a novel benchmark that departs from the paradigm of augmenting existing video datasets with text annotations and focuses on crawling and aligning available audio descriptions of mainstream movies.

421, TITLE: EvUnroll: Neuromorphic Events Based Rolling Shutter Image Correction
http://openaccess.thecvf.com//content/CVPR2022/html/Zhou_EvUnroll_Neuromorphic_Events_Based_Rolling_Shutter_Image_Correction_CVPR_2022_paper.html
AUTHORS: Xinyu Zhou, Peiqi Duan, Yi Ma, Boxin Shi
HIGHLIGHT: This paper proposes to use neuromorphic events for correcting rolling shutter (RS) images as consecutive global shutter (GS) frames.

422, TITLE: Gait Recognition in the Wild With Dense 3D Representations and a Benchmark
http://openaccess.thecvf.com//content/CVPR2022/html/Zheng_Gait_Recognition_in_the_Wild_With_Dense_3D_Representations_and_CVPR_2022_paper.html
AUTHORS: Jinkai Zheng, Xinchen Liu, Wu Liu, Lingxiao He, Chenggang Yan, Tao Mei
HIGHLIGHT: In particular, we propose a novel framework to explore the 3D Skinned Multi-Person Linear (SMPL) model of the human body for gait recognition, named SMPLGait.

423, TITLE: ArtiBoost: Boosting Articulated 3D Hand-Object Pose Estimation via Online Exploration and Synthesis
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_ArtiBoost_Boosting_Articulated_3D_Hand-Object_Pose_Estimation_via_Online_Exploration_CVPR_2022_paper.html
AUTHORS: Lixin Yang, Kailin Li, Xinyu Zhan, Jun Lv, Wenqiang Xu, Jiefeng Li, Cewu Lu
HIGHLIGHT: However, constructing both valid and diverse hand-object interactions and efficiently learning from the vast synthetic data is still challenging. To address the above issues, we propose ArtiBoost, a lightweight online data enhancement method.

424, TITLE: Temporal Context Matters: Enhancing Single Image Prediction With Disease Progression Representations
http://openaccess.thecvf.com//content/CVPR2022/html/Konwer_Temporal_Context_Matters_Enhancing_Single_Image_Prediction_With_Disease_Progression_CVPR_2022_paper.html
AUTHORS: Aishik Konwer, Xuan Xu, Joseph Bae, Chao Chen, Prateek Prasanna
HIGHLIGHT: We present a deep learning approach that leverages temporal progression information to improve clinical outcome predictions from single-timepoint images.

425, TITLE: QueryDet: Cascaded Sparse Query for Accelerating High-Resolution Small Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_QueryDet_Cascaded_Sparse_Query_for_Accelerating_High-Resolution_Small_Object_Detection_CVPR_2022_paper.html
AUTHORS: Chenhongyi Yang, Zehao Huang, Naiyan Wang
HIGHLIGHT: To get the best of two worlds, we propose QueryDet that uses a novel query mechanism to accelerate the inference speed of feature-pyramid based object detectors.

426, TITLE: IDEA-Net: Dynamic 3D Point Cloud Interpolation via Deep Embedding Alignment
http://openaccess.thecvf.com//content/CVPR2022/html/Zeng_IDEA-Net_Dynamic_3D_Point_Cloud_Interpolation_via_Deep_Embedding_Alignment_CVPR_2022_paper.html
AUTHORS: Yiming Zeng, Yue Qian, Qijian Zhang, Junhui Hou, Yixuan Yuan, Ying He
HIGHLIGHT: To tackle the challenges, we propose IDEA-Net, an end-to-end deep learning framework, which disentangles the problem under the assistance of the explicitly learned temporal consistency.

427, TITLE: UniCon: Combating Label Noise Through Uniform Selection and Contrastive Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Karim_UniCon_Combating_Label_Noise_Through_Uniform_Selection_and_Contrastive_Learning_CVPR_2022_paper.html
AUTHORS: Nazmul Karim, Mamshad Nayeem Rizve, Nazanin Rahnavard, Ajmal Mian, Mubarak Shah
HIGHLIGHT: In this work, we propose UNICON, a simple yet effective sample selection method which is robust to high label noise.

428, TITLE: Learning From All Vehicles
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Learning_From_All_Vehicles_CVPR_2022_paper.html
AUTHORS: Dian Chen, Philipp Krähenbühl
HIGHLIGHT: In this paper, we present a system to train driving policies from experiences collected not just from the ego-vehicle, but all vehicles that it observes.

429, TITLE: BEHAVE: Dataset and Method for Tracking Human Object Interactions
http://openaccess.thecvf.com//content/CVPR2022/html/Bhatnagar_BEHAVE_Dataset_and_Method_for_Tracking_Human_Object_Interactions_CVPR_2022_paper.html
AUTHORS: Bharat Lal Bhatnagar, Xianghui Xie, Ilya A. Petrov, Cristian Sminchisescu, Christian Theobalt, Gerard Pons-Moll
HIGHLIGHT: Our key insight is to predict correspondences from the human and the object to a statistical body model to obtain human-object contacts during interactions.

430, TITLE: Disentangled3D: Learning a 3D Generative Model With Disentangled Geometry and Appearance From Monocular Images
http://openaccess.thecvf.com//content/CVPR2022/html/Tewari_Disentangled3D_Learning_a_3D_Generative_Model_With_Disentangled_Geometry_and_CVPR_2022_paper.html
AUTHORS: Ayush Tewari, Mallikarjun B R, Xingang Pan, Ohad Fried, Maneesh Agrawala, Christian Theobalt
HIGHLIGHT: In this paper, we design a 3D GAN which can learn a disentangled model of objects, just from monocular observations.

431, TITLE: Revisiting Random Channel Pruning for Neural Network Compression
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Revisiting_Random_Channel_Pruning_for_Neural_Network_Compression_CVPR_2022_paper.html
AUTHORS: Yawei Li, Kamil Adamczewski, Wen Li, Shuhang Gu, Radu Timofte, Luc Van Gool
HIGHLIGHT: In this paper, we try to determine the channel configuration of the pruned models by random search.

432, TITLE: One-Bit Active Query With Contrastive Pairs
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_One-Bit_Active_Query_With_Contrastive_Pairs_CVPR_2022_paper.html
AUTHORS: Yuhang Zhang, Xiaopeng Zhang, Lingxi Xie, Jie Li, Robert C. Qiu, Hengtong Hu, Qi Tian
HIGHLIGHT: In this paper, we present a new active learning framework, which for the first time incorporates contrastive learning into recently proposed one-bit supervision.

433, TITLE:		Estimating Egocentric 3D Human Pose in the Wild With External Weak Supervision
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Estimating_Egocentric_3D_Human_Pose_in_the_Wild_With_External
_CVPR_2022_paper.html
AUTHORS:		Jian Wang, Lingjie Liu, Weipeng Xu, Kripasindhu Sarkar, Diogo Luvizon, Christian Theobalt
HIGHLIGHT:		We present a new egocentric pose estimation method, which can be trained on the new dataset with weak
external supervision.

434, TITLE:		Performance-Aware Mutual Knowledge Distillation for Improving Neural Architecture Search
http://openaccess.thecvf.com//content/CVPR2022/html/Xie_Performance-
Aware_Mutual_Knowledge_Distillation_for_Improving_Neural_Architecture_Search_CVPR_2022_paper.html
AUTHORS:		Pengtao Xie, Xuefeng Du
HIGHLIGHT:		In existing MKD methods, mutual knowledge distillation is performed between models without scrutiny: a
worse-performing model is allowed to generate knowledge to train a better-performing model, which may lead to collective failures.
To address this problem, we propose a performance-aware MKD (PAMKD) approach for NAS, where knowledge generated by model
A is allowed to train model B only if the performance of A is better than B.

435, TITLE:		Does Text Attract Attention on E-Commerce Images: A Novel Saliency Prediction Dataset and Method
http://openaccess.thecvf.com//content/CVPR2022/html/Jiang_Does_Text_Attract_Attention_on_E-
Commerce_Images_A_Novel_Saliency_CVPR_2022_paper.html
AUTHORS:		Lai Jiang, Yifei Li, Shengxi Li, Mai Xu, Se Lei, Yichen Guo, Bo Huang
HIGHLIGHT:		In this paper, we establish the first dataset of saliency e-commerce images (SalECI), which allows for learning
to predict saliency on the e-commerce images.

436, TITLE:		Topologically-Aware Deformation Fields for Single-View 3D Reconstruction
http://openaccess.thecvf.com//content/CVPR2022/html/Duggal_Topologically-Aware_Deformation_Fields_for_Single-
View_3D_Reconstruction_CVPR_2022_paper.html
AUTHORS:		Shivam Duggal, Deepak Pathak
HIGHLIGHT:		We present a new framework to learn dense 3D reconstruction and correspondence from a single 2D image.

437, TITLE:		HyperInverter: Improving StyleGAN Inversion via Hypernetwork
http://openaccess.thecvf.com//content/CVPR2022/html/Dinh_HyperInverter_Improving_StyleGAN_Inversion_via_Hypernetwork_C
VPR_2022_paper.html
AUTHORS:		Tan M. Dinh, Anh Tuan Tran, Rang Nguyen, Binh-Son Hua
HIGHLIGHT:		nfortunately, the majority of existing GAN inversion methods fail to meet at least one of the three requirements
listed below: high reconstruction quality, editability, and fast inference. We present a novel two-phase strategy in this research that fits
all requirements at the same time.

438, TITLE:		Sparse Non-Local CRF
http://openaccess.thecvf.com//content/CVPR2022/html/Veksler_Sparse_Non-Local_CRF_CVPR_2022_paper.html
AUTHORS:		Olga Veksler, Yuri Boykov
HIGHLIGHT:		We propose a new sparse non-local CRF: it has a sparse number of connections, but it has both local and non-
local ones.

439, TITLE:		Dataset Distillation by Matching Training Trajectories
http://openaccess.thecvf.com//content/CVPR2022/html/Cazenavette_Dataset_Distillation_by_Matching_Training_Trajectories_CVPR
_2022_paper.html
AUTHORS:		George Cazenavette, Tongzhou Wang, Antonio Torralba, Alexei A. Efros, Jun-Yan Zhu
HIGHLIGHT:		In this paper, we propose a new formulation that optimizes our distilled data to guide networks to a similar state
as those trained on real data across many training steps.

440, TITLE:		Towards Driving-Oriented Metric for Lane Detection Models
http://openaccess.thecvf.com//content/CVPR2022/html/Sato_Towards_Driving-
Oriented_Metric_for_Lane_Detection_Models_CVPR_2022_paper.html
AUTHORS:		Takami Sato, Qi Alfred Chen
HIGHLIGHT:		In this study, we design 2 new driving-oriented metrics for lane detection: End-to-End Lateral Deviation metric
(E2E-LD) is directly formulated based on the requirements of autonomous driving, a core task downstream of lane detection; Per-
frame Simulated Lateral Deviation metric (PSLD) is a lightweight surrogate metric of E2E-LD.

441, TITLE:		EPro-PnP: Generalized End-to-End Probabilistic Perspective-N-Points for Monocular Object Pose Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_EPro-PnP_Generalized_End-to-End_Probabilistic_Perspective-N-
Points_for_Monocular_Object_Pose_Estimation_CVPR_2022_paper.html
AUTHORS:		Hansheng Chen, Pichao Wang, Fan Wang, Wei Tian, Lu Xiong, Hao Li
HIGHLIGHT:		In this paper, we propose the EPro-PnP, a probabilistic PnP layer for general end-to-end pose estimation, which
outputs a distribution of pose on the SE(3) manifold, essentially bringing categorical Softmax to the continuous domain.

442, TITLE:		Rethinking Reconstruction Autoencoder-Based Out-of-Distribution Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Zhou_Rethinking_Reconstruction_Autoencoder-Based_Out-of-
Distribution_Detection_CVPR_2022_paper.html
AUTHORS:		Yibo Zhou

HIGHLIGHT: With desirable characteristics, reconstruction autoencoder-based methods deal with this problem by using input reconstruction error as a metric of novelty vs. normality. We formulate the essence of such approach as a quadruplet domain translation with an intrinsic bias to only query for a proxy of conditional data uncertainty.

443, TITLE: XYDeblur: Divide and Conquer for Single Image Deblurring
http://openaccess.thecvf.com//content/CVPR2022/html/Ji_XYDeblur_Divide_and_Conquer_for_Single_Image_Deblurring_CVPR_2022_paper.html
AUTHORS: Seo-Won Ji, Jeongmin Lee, Seung-Wook Kim, Jun-Pyo Hong, Seung-Jin Baek, Seung-Won Jung, Sung-Jea Ko
HIGHLIGHT: Toward an effective network architecture, we present complemental sub-solutions learning with a one-encoder-two-decoder architecture for single image deblurring.

444, TITLE: Generating Diverse and Natural 3D Human Motions From Text
http://openaccess.thecvf.com//content/CVPR2022/html/Guo_Generating_Diverse_and_Natural_3D_Human_Motions_From_Text_CVPR_2022_paper.html
AUTHORS: Chuan Guo, Shihao Zou, Xinxin Zuo, Sen Wang, Wei Ji, Xingyu Li, Li Cheng
HIGHLIGHT: Instead of directly engaging with pose sequences, we propose motion snippet code as our internal motion representation, which captures local semantic motion contexts and is empirically shown to facilitate the generation of plausible motions faithful to the input text.

445, TITLE: E-CIR: Event-Enhanced Continuous Intensity Recovery
http://openaccess.thecvf.com//content/CVPR2022/html/Song_E-CIR_Event-Enhanced_Continuous_Intensity_Recovery_CVPR_2022_paper.html
AUTHORS: Chen Song, Qixing Huang, Chandrajit Bajaj
HIGHLIGHT: This paper presents E-CIR, which converts a blurry image into a sharp video represented as a parametric function from time to intensity.

446, TITLE: Towards Robust Rain Removal Against Adversarial Attacks: A Comprehensive Benchmark Analysis and Beyond
http://openaccess.thecvf.com//content/CVPR2022/html/Yu_Towards_Robust_Rain_Removal_Against_Adversarial_Attacks_A_Comprehensive_Benchmark_CVPR_2022_paper.html
AUTHORS: Yi Yu, Wenhan Yang, Yap-Peng Tan, Alex C. Kot
HIGHLIGHT: A systematic evaluation of key modules in existing methods is performed in terms of their robustness against adversarial attacks. From the insights of our analysis, we construct a more robust deraining method by integrating these effective modules.

447, TITLE: STCrowd: A Multimodal Dataset for Pedestrian Perception in Crowded Scenes
http://openaccess.thecvf.com//content/CVPR2022/html/Cong_STCrowd_A_Multimodal_Dataset_for_Pedestrian_Perception_in_Crowded_Scenes_CVPR_2022_paper.html
AUTHORS: Peishan Cong, Xinge Zhu, Feng Qiao, Yiming Ren, Xidong Peng, Yuenan Hou, Lan Xu, Ruigang Yang, Dinesh Manocha, Yuexin Ma
HIGHLIGHT: To better evaluate pedestrian perception algorithms in crowded scenarios, we introduce a large-scale multimodal dataset, STCrowd.

448, TITLE: Deep Decomposition for Stochastic Normal-Abnormal Transport
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Deep_Decomposition_for_Stochastic_Normal-Abnormal_Transport_CVPR_2022_paper.html
AUTHORS: Peirong Liu, Yueh Lee, Stephen Aylward, Marc Niethammer
HIGHLIGHT: We develop a machine learning model, D^2-SONATA, built upon a stochastic advection-diffusion equation, which predicts the velocity and diffusion fields that drive 2D/3D image time-series of transport.

449, TITLE: Global Context With Discrete Diffusion in Vector Quantised Modelling for Image Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Hu_Global_Context_With_Discrete_Diffusion_in_Vector_Quantised_Modelling_for_CVPR_2022_paper.html
AUTHORS: Minghui Hu, Yujie Wang, Tat-Jen Cham, Jianfei Yang, P.N. Suganthan
HIGHLIGHT: We show that with the help of a content-rich discrete visual codebook from VQ-VAE, the discrete diffusion model can also generate high fidelity images with global context, which compensates for the deficiency of the classical autoregressive model along pixel space.

450, TITLE: Symmetry and Uncertainty-Aware Object SLAM for 6DoF Object Pose Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Merrill_Symmetry_and_Uncertainty-Aware_Object_SLAM_for_6DoF_Object_Pose_Estimation_CVPR_2022_paper.html
AUTHORS: Nathaniel Merrill, Yuliang Guo, Xingxing Zuo, Xinyu Huang, Stefan Leutenegger, Xi Peng, Liu Ren, Guoquan Huang
HIGHLIGHT: We propose a keypoint-based object-level SLAM framework that can provide globally consistent 6DoF pose estimates for symmetric and asymmetric objects alike.

451, TITLE: AziNorm: Exploiting the Radial Symmetry of Point Cloud for Azimuth-Normalized 3D Perception
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_AziNorm_Exploiting_the_Radial_Symmetry_of_Point_Cloud_for_Azimuth-Normalized_CVPR_2022_paper.html
AUTHORS: Shaoyu Chen, Xinggang Wang, Tianheng Cheng, Wenqiang Zhang, Qian Zhang, Chang Huang, Wenyu Liu

HIGHLIGHT: Point cloud, the most important data format for 3D environmental perception, is naturally endowed with strong radial symmetry. In this work, we exploit this radial symmetry via a divide-and-conquer strategy to boost 3D perception performance and ease optimization.

452, TITLE: Towards Multimodal Depth Estimation From Light Fields
http://openaccess.thecvf.com//content/CVPR2022/html/Leistner_Towards_Multimodal_Depth_Estimation_From_Light_Fields_CVPR_2022_paper.html
AUTHORS: Titus Leistner, Radek Mackowiak, Lynton Ardizzone, Ullrich Köthe, Carsten Rother
HIGHLIGHT: This allows us to supervise the multimodal depth prediction and also validate all methods by measuring the KL divergence of the predicted posteriors. With our thorough analysis and novel dataset, we aim to start a new line of depth estimation research that overcomes some of the long-standing limitations of this field.

453, TITLE: Learning To Recognize Procedural Activities With Distant Supervision
http://openaccess.thecvf.com//content/CVPR2022/html/Lin_Learning_To_Recognize_Procedural_Activities_With_Distant_Supervision_CVPR_2022_paper.html
AUTHORS: Xudong Lin, Fabio Petroni, Gedas Bertasius, Marcus Rohrbach, Shih-Fu Chang, Lorenzo Torresani
HIGHLIGHT: In this paper we consider the problem of classifying fine-grained, multi-step activities (e.g., cooking different recipes, making disparate home improvements, creating various forms of arts and crafts) from long videos spanning up to several minutes.

454, TITLE: Multimodal Material Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Liang_Multimodal_Material_Segmentation_CVPR_2022_paper.html
AUTHORS: Yupeng Liang, Ryosuke Wakaki, Shohei Nobuhara, Ko Nishino
HIGHLIGHT: We introduce the MCubeS dataset (from MultiModal Material Segmentation) which contains 500 sets of multimodal images capturing 42 street scenes.

455, TITLE: Multi-Frame Self-Supervised Depth With Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Guizilini_Multi-Frame_Self-Supervised_Depth_With_Transformers_CVPR_2022_paper.html
AUTHORS: Vitor Guizilini, RareÈ™ AmbruÈ™, Dian Chen, Sergey Zakharov, Adrien Gaidon
HIGHLIGHT: In this paper we revisit feature matching for self-supervised monocular depth estimation, and propose a novel transformer architecture for cost volume generation.

456, TITLE: Weakly Supervised Rotation-Invariant Aerial Object Detection Network
http://openaccess.thecvf.com//content/CVPR2022/html/Feng_Weakly_Supervised_Rotation-Invariant_Aerial_Object_Detection_Network_CVPR_2022_paper.html
AUTHORS: Xiaoxu Feng, Xiwen Yao, Gong Cheng, Junwei Han
HIGHLIGHT: Meanwhile, current solutions have been prone to fall into the issue with unstable detectors, as they ignore lower-scored instances and may regard them as backgrounds. To address these issues, in this paper, we construct a novel end-to-end weakly supervised Rotation-Invariant aerial object detection Network (RINet).

457, TITLE: Modeling Motion With Multi-Modal Features for Text-Based Video Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Zhao_Modeling_Motion_With_Multi-Modal_Features_for_Text-Based_Video_Segmentation_CVPR_2022_paper.html
AUTHORS: Wangbo Zhao, Kai Wang, Xiangxiang Chu, Fuzhao Xue, Xinchao Wang, Yang You
HIGHLIGHT: In this paper, we design a method to fuse and align appearance, motion, and linguistic features to achieve accurate segmentation.

458, TITLE: Surface Reconstruction From Point Clouds by Learning Predictive Context Priors
http://openaccess.thecvf.com//content/CVPR2022/html/Ma_Surface_Reconstruction_From_Point_Clouds_by_Learning_Predictive_Context_Priors_CVPR_2022_paper.html
AUTHORS: Baorui Ma, Yu-Shen Liu, Matthias Zwicker, Zhizhong Han
HIGHLIGHT: However, this requires the local context prior to generalize to a wide variety of unseen target regions, which is hard to achieve. To resolve this issue, we introduce Predictive Context Priors by learning Predictive Queries for each specific point cloud at inference time.

459, TITLE: Deformable Video Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Deformable_Video_Transformer_CVPR_2022_paper.html
AUTHORS: Jue Wang, Lorenzo Torresani
HIGHLIGHT: In this paper, we introduce the Deformable Video Transformer (DVT), which dynamically predicts a small subset of video patches to attend for each query location based on motion information, thus allowing the model to decide where to look in the video based on correspondences across frames.

460, TITLE: Self-Supervised Keypoint Discovery in Behavioral Videos
http://openaccess.thecvf.com//content/CVPR2022/html/Sun_Self-Supervised_Keypoint_Discovery_in_Behavioral_Videos_CVPR_2022_paper.html
AUTHORS: Jennifer J. Sun, Serim Ryou, Roni H. Goldshmid, Brandon Weissbourd, John O. Dabiri, David J. Anderson, Ann Kennedy, Yisong Yue, Pietro Perona
HIGHLIGHT: We propose a method for learning the posture and structure of agents from unlabelled behavioral videos.

461, TITLE:	IRISformer: Dense Vision Transformers for Single-Image Inverse Rendering in Indoor Scenes
http://openaccess.thecvf.com//content/CVPR2022/html/Zhu_IRISformer_Dense_Vision_Transformers_for_Single-Image_Inverse_Rendering_in_Indoor_CVPR_2022_paper.html
AUTHORS:	Rui Zhu, Zhengqin Li, Janarbek Matai, Fatih Porikli, Manmohan Chandraker
HIGHLIGHT:	In this work, our intuition is that the long-range attention learned by transformer architectures is ideally suited to solve longstanding challenges in single-image inverse rendering.

462, TITLE:	DynamicEarthNet: Daily Multi-Spectral Satellite Dataset for Semantic Change Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Toker_DynamicEarthNet_Daily_Multi-Spectral_Satellite_Dataset_for_Semantic_Change_Segmentation_CVPR_2022_paper.html
AUTHORS:	Aysim Toker, Lukas Kondmann, Mark Weber, Marvin Eisenberger, Andrés Camero, Jingliang Hu, Ariadna Pregel Hoderlein, Çaĝlar Åženaras, Timothy Davis, Daniel Cremers, Giovanni Marchisio, Xiao Xiang Zhu, Laura Leal-Taixé
HIGHLIGHT:	To that end, we propose the DynamicEarthNet dataset that consists of daily, multi-spectral satellite observations of 75 selected areas of interest distributed over the globe with imagery from Planet Labs.

463, TITLE:	Connecting the Complementary-View Videos: Joint Camera Identification and Subject Association
http://openaccess.thecvf.com//content/CVPR2022/html/Han_Connecting_the_Complementary-View_Videos_Joint_Camera_Identification_and_Subject_Association_CVPR_2022_paper.html
AUTHORS:	Ruize Han, Yiyang Gan, Jiacheng Li, Feifan Wang, Wei Feng, Song Wang
HIGHLIGHT:	In this paper, we develop a new approach that can simultaneously handle three tasks: i) localizing the side-view camera in the top view; ii) estimating the view direction of the side-view camera; iii) detecting and associating the same subjects on the ground across the complementary views.

464, TITLE:	End-to-End Trajectory Distribution Prediction Based on Occupancy Grid Maps
http://openaccess.thecvf.com//content/CVPR2022/html/Guo_End-to-End_Trajectory_Distribution_Prediction_Based_on_Occupancy_Grid_Maps_CVPR_2022_paper.html
AUTHORS:	Ke Guo, Wenxi Liu, Jia Pan
HIGHLIGHT:	In this paper, we aim to forecast a future trajectory distribution of a moving agent in the real world, given the social scene images and historical trajectories.

465, TITLE:	Fast, Accurate and Memory-Efficient Partial Permutation Synchronization
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Fast_Accurate_and_Memory-Efficient_Partial_Permutation_Synchronization_CVPR_2022_paper.html
AUTHORS:	Shaohan Li, Yunpeng Shi, Gilad Lerman
HIGHLIGHT:	Here we overcome the restriction of CEMP to compact groups and propose an improved algorithm, CEMP-Partial, for estimating the corruption levels of the observed partial permutations.

466, TITLE:	Quantization-Aware Deep Optics for Diffractive Snapshot Hyperspectral Imaging
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Quantization-Aware_Deep_Optics_for_Diffractive_Snapshot_Hyperspectral_Imaging_CVPR_2022_paper.html
AUTHORS:	Lingen Li, Lizhi Wang, Weitao Song, Lei Zhang, Zhiwei Xiong, Hua Huang
HIGHLIGHT:	In this paper, we propose the quantization-aware deep optics for diffractive snapshot hyperspectral imaging.

467, TITLE:	Weakly Supervised Temporal Action Localization via Representative Snippet Knowledge Propagation
http://openaccess.thecvf.com//content/CVPR2022/html/Huang_Weakly_Supervised_Temporal_Action_Localization_via_Representative_Snippet_Knowledge_Propagation_CVPR_2022_paper.html
AUTHORS:	Linjiang Huang, Liang Wang, Hongsheng Li
HIGHLIGHT:	Many existing methods seek to generate pseudo labels for bridging the discrepancy between classification and localization, but usually only make use of limited contextual information for pseudo label generation. To alleviate this problem, we propose a representative snippet summarization and propagation framework.

468, TITLE:	Parametric Scattering Networks
http://openaccess.thecvf.com//content/CVPR2022/html/Gauthier_Parametric_Scattering_Networks_CVPR_2022_paper.html
AUTHORS:	Shanel Gauthier, Benjamin Thérien, Laurent Alsène-Racicot, Muawiz Chaudhary, Irina Rish, Eugene Belilovsky, Michael Eickenberg, Guy Wolf
HIGHLIGHT:	The wavelet filters used in the scattering transform are typically selected to create a tight frame via a parameterized mother wavelet. In this work, we investigate whether this standard wavelet filterbank construction is optimal.

469, TITLE:	SketchEdit: Mask-Free Local Image Manipulation With Partial Sketches
http://openaccess.thecvf.com//content/CVPR2022/html/Zeng_SketchEdit_Mask-Free_Local_Image_Manipulation_With_Partial_Sketches_CVPR_2022_paper.html
AUTHORS:	Yu Zeng, Zhe Lin, Vishal M. Patel
HIGHLIGHT:	Although this setup simplifies data preparation and model design, it complicates user interaction and discards useful information in masked regions. To this end, we propose a new framework for sketch-based image manipulation that only requires sketch inputs from users and utilizes the entire original image.

470, TITLE:	ScaleNet: A Shallow Architecture for Scale Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Barroso-Laguna_ScaleNet_A_Shallow_Architecture_for_Scale_Estimation_CVPR_2022_paper.html

AUTHORS: Axel Barroso-Laguna, Yurun Tian, Krystian Mikolajczyk
HIGHLIGHT: In this paper, we address the problem of estimating scale factors between images.

471, TITLE: E2EC: An End-to-End Contour-Based Method for High-Quality High-Speed Instance Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_E2EC_An_End-to-End_Contour-Based_Method_for_High-Quality_High-Speed_Instance_Segmentation_CVPR_2022_paper.html
AUTHORS: Tao Zhang, Shiqing Wei, Shunping Ji
HIGHLIGHT: In this paper, we introduce a novel contour-based method, named E2EC, for high-quality instance segmentation.

472, TITLE: Bounded Adversarial Attack on Deep Content Features
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_Bounded_Adversarial_Attack_on_Deep_Content_Features_CVPR_2022_paper.html
AUTHORS: Qiuling Xu, Guanhong Tao, Xiangyu Zhang
HIGHLIGHT: We propose a novel adversarial attack targeting content features in some deep layer, that is, individual neurons in the layer.

473, TITLE: BatchFormer: Learning To Explore Sample Relationships for Robust Representation Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Hou_BatchFormer_Learning_To_Explore_Sample_Relationships_for_Robust_Representation_Learning_CVPR_2022_paper.html
AUTHORS: Zhi Hou, Baosheng Yu, Dacheng Tao
HIGHLIGHT: To address the above-mentioned issues, a variety of methods have been devised to explore the sample relationships in a vanilla way (i.e., from the perspectives of either the input or the loss function), failing to explore the internal structure of deep neural networks for learning with sample relationships. Inspired by this, we propose to enable deep neural networks themselves with the ability to learn the sample relationships from each mini-batch.

474, TITLE: Self-Supervised Image-Specific Prototype Exploration for Weakly Supervised Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Self-Supervised_Image-Specific_Prototype_Exploration_for_Weakly_Supervised_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Qi Chen, Lingxiao Yang, Jian-Huang Lai, Xiaohua Xie
HIGHLIGHT: However, the classifier focuses only on the discriminative regions while ignoring other useful information in each image, resulting in incomplete localization maps. To address this issue, we propose a Self-supervised Image-specific Prototype Exploration (SIPE) that consists of an Image-specific Prototype Exploration (IPE) and a General-Specific Consistency (GSC) loss.

475, TITLE: CAD: Co-Adapting Discriminative Features for Improved Few-Shot Classification
http://openaccess.thecvf.com//content/CVPR2022/html/Chikontwe_CAD_Co-Adapting_Discriminative_Features_for_Improved_Few-Shot_Classification_CVPR_2022_paper.html
AUTHORS: Philip Chikontwe, Soopil Kim, Sang Hyun Park
HIGHLIGHT: In this paper, inspired by the transformer style self-attention mechanism, we propose a strategy to cross-attend and re-weight discriminative features for few-shot classification.

476, TITLE: Fingerprinting Deep Neural Networks Globally via Universal Adversarial Perturbations
http://openaccess.thecvf.com//content/CVPR2022/html/Peng_Fingerprinting_Deep_Neural_Networks_Globally_via_Universal_Adversarial_Perturbations_CVPR_2022_paper.html
AUTHORS: Zirui Peng, Shaofeng Li, Guoxing Chen, Cheng Zhang, Haojin Zhu, Minhui Xue
HIGHLIGHT: In this paper, we propose a novel and practical mechanism which enables the service provider to verify whether a suspect model is stolen from the victim model via model extraction attacks.

477, TITLE: Learning Multi-View Aggregation in the Wild for Large-Scale 3D Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Robert_Learning_Multi-View_Aggregation_in_the_Wild_for_Large-Scale_3D_Semantic_CVPR_2022_paper.html
AUTHORS: Damien Robert, Bruno Vallet, Loic Landrieu
HIGHLIGHT: In contrast, we propose an end-to-end trainable multi-view aggregation model leveraging the viewing conditions of 3D points to merge features from images taken at arbitrary positions.

478, TITLE: ManiTrans: Entity-Level Text-Guided Image Manipulation via Token-Wise Semantic Alignment and Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_ManiTrans_Entity-Level_Text-Guided_Image_Manipulation_via_Token-Wise_Semantic_Alignment_and_CVPR_2022_paper.html
AUTHORS: Jianan Wang, Guansong Lu, Hang Xu, Zhenguo Li, Chunjing Xu, Yanwei Fu
HIGHLIGHT: In this work, we study a novel task on text-guided image manipulation on the entity level in the real world.

479, TITLE: Improving Video Model Transfer With Dynamic Representation Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Improving_Video_Model_Transfer_With_Dynamic_Representation_Learning_CVPR_2022_paper.html
AUTHORS: Yi Li, Nuno Vasconcelos
HIGHLIGHT: In this work, the problem of dynamic representation learning (DRL) is studied.

480, TITLE: PIE-Net: Photometric Invariant Edge Guided Network for Intrinsic Image Decomposition
http://openaccess.thecvf.com//content/CVPR2022/html/Das_PIE-Net_Photometric_Invariant_Edge_Guided_Network_for_Intrinsic_Image_Decomposition_CVPR_2022_paper.html

AUTHORS: Partha Das, Sezer Karaoglu, Theo Gevers
HIGHLIGHT: These methods can be negatively influenced by strong illumination conditions causing shading-reflectance leakages. Therefore, in this paper, an end-to-end edge-driven hybrid CNN approach is proposed for intrinsic image decomposition.

481, TITLE: Clothes-Changing Person Re-Identification With RGB Modality Only
http://openaccess.thecvf.com//content/CVPR2022/html/Gu_Clothes-Changing_Person_Re-Identification_With_RGB_Modality_Only_CVPR_2022_paper.html
AUTHORS: Xinqian Gu, Hong Chang, Bingpeng Ma, Shutao Bai, Shiguang Shan, Xilin Chen
HIGHLIGHT: In this paper, we propose a Clothes-based Adversarial Loss (CAL) to mine clothes-irrelevant features from the original RGB images by penalizing the predictive power of re-id model w.r.t. clothes.

482, TITLE: Chitransformer: Towards Reliable Stereo From Cues
http://openaccess.thecvf.com//content/CVPR2022/html/Su_Chitransformer_Towards_Reliable_Stereo_From_Cues_CVPR_2022_paper.html
AUTHORS: Qing Su, Shihao Ji
HIGHLIGHT: While single image depth estimation is spared from these challenges and can achieve satisfactory results with the extracted monocular cues, the lack of stereoscopic relationship renders the monocular prediction less reliable on its own, especially in highly dynamic or cluttered environments. To address these issues in both scenarios, we present an optic-chiasm-inspired self-supervised binocular depth estimation method, wherein a vision transformer (ViT) with gated positional cross-attention (GPCA) layers is designed to enable feature-sensitive pattern retrieval between views while retaining the extensive context information aggregated through self-attentions.

483, TITLE: Robust Image Forgery Detection Over Online Social Network Shared Images
http://openaccess.thecvf.com//content/CVPR2022/html/Wu_Robust_Image_Forgery_Detection_Over_Online_Social_Network_Shared_Images_CVPR_2022_paper.html
AUTHORS: Haiwei Wu, Jiantao Zhou, Jinyu Tian, Jun Liu
HIGHLIGHT: To fight against the OSN-shared forgeries, in this work, a novel robust training scheme is proposed.

484, TITLE: QS-Attn: Query-Selected Attention for Contrastive Learning in I2I Translation
http://openaccess.thecvf.com//content/CVPR2022/html/Hu_QS-Attn_Query-Selected_Attention_for_Contrastive_Learning_in_I2I_Translation_CVPR_2022_paper.html
AUTHORS: Xueqi Hu, Xinyue Zhou, Qiusheng Huang, Zhengyi Shi, Li Sun, Qingli Li
HIGHLIGHT: Moreover, the feature itself does not reflect the relation with others. This paper deals with these problems by intentionally selecting significant anchor points for contrastive learning.

485, TITLE: Physically Disentangled Intra- and Inter-Domain Adaptation for Varicolored Haze Removal
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Physically_Disentangled_Intra-_and_Inter-Domain_Adaptation_for_Varicolored_Haze_Removal_CVPR_2022_paper.html
AUTHORS: Yi Li, Yi Chang, Yan Gao, Changfeng Yu, Luxin Yan
HIGHLIGHT: In this work, our goal is to handle the new task: real-world varicolored haze removal.

486, TITLE: Modality-Agnostic Learning for Radar-Lidar Fusion in Vehicle Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Modality-Agnostic_Learning_for_Radar-Lidar_Fusion_in_Vehicle_Detection_CVPR_2022_paper.html
AUTHORS: Yu-Jhe Li, Jinhyung Park, Matthew O'Toole, Kris Kitani
HIGHLIGHT: If one of the sensors is unavailable or missing, the model may fail catastrophically. To mitigate this problem, we propose the Self-Training Multimodal Vehicle Detection Network (ST-MVDNet) which leverages a Teacher-Student mutual learning framework and a simulated sensor noise model used in strong data augmentation for Lidar and Radar.

487, TITLE: A Re-Balancing Strategy for Class-Imbalanced Classification Based on Instance Difficulty
http://openaccess.thecvf.com//content/CVPR2022/html/Yu_A_Re-Balancing_Strategy_for_Class-Imbalanced_Classification_Based_on_Instance_Difficulty_CVPR_2022_paper.html
AUTHORS: Sihao Yu, Jiafeng Guo, Ruqing Zhang, Yixing Fan, Zizhen Wang, Xueqi Cheng
HIGHLIGHT: By reducing the weights of the majority classes, such instances would become more difficult to learn and hurt the overall performance consequently. To tackle this problem, we propose a novel instance-level re-balancing strategy, which dynamically adjusts the sampling probabilities of instances according to the instance difficulty.

488, TITLE: Representation Compensation Networks for Continual Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Representation_Compensation_Networks_for_Continual_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Chang-Bin Zhang, Jia-Wen Xiao, Xialei Liu, Ying-Cong Chen, Ming-Ming Cheng
HIGHLIGHT: In this work, we study the continual semantic segmentation problem, where the deep neural networks are required to incorporate new classes continually without catastrophic forgetting.

489, TITLE: Adaptive Gating for Single-Photon 3D Imaging
http://openaccess.thecvf.com//content/CVPR2022/html/Po_Adaptive_Gating_for_Single-Photon_3D_Imaging_CVPR_2022_paper.html
AUTHORS: Ryan Po, Adithya Pediredla, Ioannis Gkioulekas
HIGHLIGHT: We propose an adaptive gating scheme built upon Thompson sampling.

490, TITLE: Tracking People by Predicting 3D Appearance, Location and Pose
http://openaccess.thecvf.com//content/CVPR2022/html/Rajasegaran_Tracking_People_by_Predicting_3D_Appearance_Location_and
_Pose_CVPR_2022_paper.html
AUTHORS: Jathushan Rajasegaran, Georgios Pavlakos, Angjoo Kanazawa, Jitendra Malik
HIGHLIGHT: We present an approach for tracking people in monocular videos by predicting their future 3D representations.

491, TITLE: Text2Mesh: Text-Driven Neural Stylization for Meshes
http://openaccess.thecvf.com//content/CVPR2022/html/Michel_Text2Mesh_Text-
Driven_Neural_Stylization_for_Meshes_CVPR_2022_paper.html
AUTHORS: Oscar Michel, Roi Bar-On, Richard Liu, Sagie Benaim, Rana Hanocka
HIGHLIGHT: In this work, we develop intuitive controls for editing the style of 3D objects.

492, TITLE: Learning To Solve Hard Minimal Problems
http://openaccess.thecvf.com//content/CVPR2022/html/Hruby_Learning_To_Solve_Hard_Minimal_Problems_CVPR_2022_paper.ht
ml
AUTHORS: Petr Hruby, Timothy Duff, Anton Leykin, Tomas Pajdla
HIGHLIGHT: We present an approach to solving hard geometric optimization problems in the RANSAC framework.

493, TITLE: H4D: Human 4D Modeling by Learning Neural Compositional Representation
http://openaccess.thecvf.com//content/CVPR2022/html/Jiang_H4D_Human_4D_Modeling_by_Learning_Neural_Compositional_Rep
resentation_CVPR_2022_paper.html
AUTHORS: Boyan Jiang, Yinda Zhang, Xingkui Wei, Xiangyang Xue, Yanwei Fu
HIGHLIGHT: This work presents a novel framework that can effectively learn a compact and compositional representation for
dynamic human by exploiting the human body prior from the widely used SMPL parametric model.

494, TITLE: FWD: Real-Time Novel View Synthesis With Forward Warping and Depth
http://openaccess.thecvf.com//content/CVPR2022/html/Cao_FWD_Real-
Time_Novel_View_Synthesis_With_Forward_Warping_and_Depth_CVPR_2022_paper.html
AUTHORS: Ang Cao, Chris Rockwell, Justin Johnson
HIGHLIGHT: In our paper, we propose a generalizable NVS method with sparse inputs, called \FWDds, which gives high-
quality synthesis in real-time.

495, TITLE: Non-Generative Generalized Zero-Shot Learning via Task-Correlated Disentanglement and Controllable
Samples Synthesis
http://openaccess.thecvf.com//content/CVPR2022/html/Feng_Non-Generative_Generalized_Zero-Shot_Learning_via_Task-
Correlated_Disentanglement_and_Controllable_Samples_CVPR_2022_paper.html
AUTHORS: Yaogong Feng, Xiaowen Huang, Pengbo Yang, Jian Yu, Jitao Sang
HIGHLIGHT: In this paper, we propose a non-generative model to address these problems correspondingly in two modules:
(1) Task-correlated feature disentanglement, to exclude the task-correlated features from task-independent ones by adversarial
learning of domain adaption towards reasonable synthesis; (2) Controllable pseudo sample synthesis, to synthesize edge-pseudo and
center-pseudo samples with certain characteristics towards more diversity generated and intuitive transfer.

496, TITLE: C-CAM: Causal CAM for Weakly Supervised Semantic Segmentation on Medical Image
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_C-
CAM_Causal_CAM_for_Weakly_Supervised_Semantic_Segmentation_on_Medical_CVPR_2022_paper.html
AUTHORS: Zhang Chen, Zhiqiang Tian, Jihua Zhu, Ce Li, Shaoyi Du
HIGHLIGHT: In this paper, we find that there are mainly two challenges of medical images in WSSS: i) the boundary of
object foreground and background is not clear; ii) the co-occurrence phenomenon is very severe in training stage. We thus propose a
Causal CAM (C-CAM) method to overcome the above challenges.

497, TITLE: Leveraging Real Talking Faces via Self-Supervision for Robust Forgery Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Haliassos_Leveraging_Real_Talking_Faces_via_Self-
Supervision_for_Robust_Forgery_Detection_CVPR_2022_paper.html
AUTHORS: Alexandros Haliassos, Rodrigo Mira, Stavros Petridis, Maja Pantic
HIGHLIGHT: One of the most pressing challenges for the detection of face-manipulated videos is generalising to forgery
methods not seen during training while remaining effective under common corruptions such as compression. In this paper, we
examine whether we can tackle this issue by harnessing videos of real talking faces, which contain rich information on natural facial
appearance and behaviour and are readily available in large quantities online.

498, TITLE: Forward Compatible Few-Shot Class-Incremental Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Zhou_Forward_Compatible_Few-Shot_Class-
Incremental_Learning_CVPR_2022_paper.html
AUTHORS: Da-Wei Zhou, Fu-Yun Wang, Han-Jia Ye, Liang Ma, Shiliang Pu, De-Chuan Zhan
HIGHLIGHT: By contrast, we suggest learning prospectively to prepare for future updates, and propose ForwArd Compatible
Training (FACT) for FSCIL.

499, TITLE: BaLeNAS: Differentiable Architecture Search via the Bayesian Learning Rule
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_BaLeNAS_Differentiable_Architecture_Search_via_the_Bayesian_Le
arning_Rule_CVPR_2022_paper.html
AUTHORS: Miao Zhang, Shirui Pan, Xiaojun Chang, Steven Su, Jilin Hu, Gholamreza (Reza) Haffari, Bin Yang

HIGHLIGHT: Rather than directly optimizing the architecture parameters, this paper formulates the neural architecture search as a distribution learning problem through relaxing the architecture weights into Gaussian distributions.

500, TITLE: Cannot See the Forest for the Trees: Aggregating Multiple Viewpoints To Better Classify Objects in Videos
http://openaccess.thecvf.com//content/CVPR2022/html/Hwang_Cannot_See_the_Forest_for_the_Trees_Aggregating_Multiple_Viewpoints_CVPR_2022_paper.html
AUTHORS: Sukjun Hwang, Miran Heo, Seoung Wug Oh, Seon Joo Kim
HIGHLIGHT: In this paper, we present a set classifier that improves accuracy of classifying tracklets by aggregating information from multiple viewpoints contained in a tracklet.

501, TITLE: Learning Canonical F-Correlation Projection for Compact Multiview Representation
http://openaccess.thecvf.com//content/CVPR2022/html/Yuan_Learning_Canonical_F-Correlation_Projection_for_Compact_Multiview_Representation_CVPR_2022_paper.html
AUTHORS: Yun-Hao Yuan, Jin Li, Yun Li, Jipeng Qiang, Yi Zhu, Xiaobo Shen, Jianping Gou
HIGHLIGHT: In this paper, we address the preceding problem and propose a novel canonical F-correlation framework by exploring and exploiting the nonlinear relationship between different features.

502, TITLE: DIFNet: Boosting Visual Information Flow for Image Captioning
http://openaccess.thecvf.com//content/CVPR2022/html/Wu_DIFNet_Boosting_Visual_Information_Flow_for_Image_Captioning_CVPR_2022_paper.html
AUTHORS: Mingrui Wu, Xuying Zhang, Xiaoshuai Sun, Yiyi Zhou, Chao Chen, Jiaxin Gu, Xing Sun, Rongrong Ji
HIGHLIGHT: However, for most cases, the partially generated sentence may dominate the target word prediction due to the insufficiency of visual information, making the generated descriptions irrelevant to the content of the given image. In this paper, we propose a Dual Information Flow Network (DIFNet) to address this issue, which takes segmentation feature as another visual information source to enhance the contribution of visual information for prediction.

503, TITLE: Weakly Supervised Object Localization As Domain Adaption
http://openaccess.thecvf.com//content/CVPR2022/html/Zhu_Weakly_Supervised_Object_Localization_As_Domain_Adaption_CVPR_2022_paper.html
AUTHORS: Lei Zhu, Qi She, Qian Chen, Yunfei You, Boyu Wang, Yanye Lu
HIGHLIGHT: However, the MIL mechanism makes CAM only activate discriminative object parts rather than the whole object, weakening its performance for localizing objects. To avoid this problem, this work provides a novel perspective that models WSOL as a domain adaption (DA) task, where the score estimator trained on the source/image domain is tested on the target/pixel domain to locate objects.

504, TITLE: Tencent-MVSE: A Large-Scale Benchmark Dataset for Multi-Modal Video Similarity Evaluation
http://openaccess.thecvf.com//content/CVPR2022/html/Zeng_Tencent-MVSE_A_Large-Scale_Benchmark_Dataset_for_Multi-Modal_Video_Similarity_Evaluation_CVPR_2022_paper.html
AUTHORS: Zhaoyang Zeng, Yongsheng Luo, Zhenhua Liu, Fengyun Rao, Dian Li, Weidong Guo, Zhen Wen
HIGHLIGHT: In this paper, we propose the Tencent-MVSE dataset, which is the first benchmark dataset for the multi-modal video similarity evaluation task.

505, TITLE: Dynamic Prototype Convolution Network for Few-Shot Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Dynamic_Prototype_Convolution_Network_for_Few-Shot_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Jie Liu, Yanqi Bao, Guo-Sen Xie, Huan Xiong, Jan-Jakob Sonke, Efstratios Gavves
HIGHLIGHT: However, these interaction approaches usually cannot well capture the intrinsic object details in the query images that are widely encountered in FSS, e.g., if the query object to be segmented has holes and slots, inaccurate segmentation almost always happens. To this end, we propose a dynamic prototype convolution network (DPCN) to fully capture the aforementioned intrinsic details for accurate FSS.

506, TITLE: Deep Orientation-Aware Functional Maps: Tackling Symmetry Issues in Shape Matching
http://openaccess.thecvf.com//content/CVPR2022/html/Donati_Deep_Orientation-Aware_Functional_Maps_Tackling_Symmetry_Issues_in_Shape_Matching_CVPR_2022_paper.html
AUTHORS: Nicolas Donati, Etienne Corman, Maks Ovsjanikov
HIGHLIGHT: Meanwhile, recent advances in the functional map framework allow to enforce orientation preservation using a functional representation for tangent vector field transfer, through so-called complex functional maps. Using this representation, we propose a new deep learning approach to learn orientation-aware features in a fully unsupervised setting.

507, TITLE: Tree Energy Loss: Towards Sparsely Annotated Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Liang_Tree_Energy_Loss_Towards_Sparsely_Annotated_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Zhiyuan Liang, Tiancai Wang, Xiangyu Zhang, Jian Sun, Jianbing Shen
HIGHLIGHT: In this paper, we propose a novel tree energy loss for SASS by providing semantic guidance for unlabeled pixels.

508, TITLE: Mr.BiQ: Post-Training Non-Uniform Quantization Based on Minimizing the Reconstruction Error
http://openaccess.thecvf.com//content/CVPR2022/html/Jeon_Mr.BiQ_Post-Training_Non-Uniform_Quantization_Based_on_Minimizing_the_Reconstruction_Error_CVPR_2022_paper.html
AUTHORS: Yongkweon Jeon, Chungman Lee, Eulrang Cho, Yeonju Ro

HIGHLIGHT: We thus propose a new post-training non-uniform quantization method, called Mr.BiQ, allowing low bit-width quantization even on Transformer models.

509, TITLE: MatteFormer: Transformer-Based Image Matting via Prior-Tokens
http://openaccess.thecvf.com//content/CVPR2022/html/Park_MatteFormer_Transformer-Based_Image_Matting_via_Prior-Tokens_CVPR_2022_paper.html
AUTHORS: GyuTae Park, SungJoon Son, JaeYoung Yoo, SeHo Kim, Nojun Kwak
HIGHLIGHT: In this paper, we propose a transformer-based image matting model called MatteFormer, which takes full advantage of trimap information in the transformer block.

510, TITLE: Video Shadow Detection via Spatio-Temporal Interpolation Consistency Training
http://openaccess.thecvf.com//content/CVPR2022/html/Lu_Video_Shadow_Detection_via_Spatio-Temporal_Interpolation_Consistency_Training_CVPR_2022_paper.html
AUTHORS: Xiao Lu, Yihong Cao, Sheng Liu, Chengjiang Long, Zipei Chen, Xuanyu Zhou, Yimin Yang, Chunxia Xiao
HIGHLIGHT: Using a model trained on labeled images to the video frames directly may lead to high generalization error and temporal inconsistent results. In this paper, we address these challenges by proposing a Spatio-Temporal Interpolation Consistency Training (STICT) framework to rationally feed the unlabeled video frames together with the labeled images into an image shadow detection network training.

511, TITLE: Ranking Distance Calibration for Cross-Domain Few-Shot Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Ranking_Distance_Calibration_for_Cross-Domain_Few-Shot_Learning_CVPR_2022_paper.html
AUTHORS: Pan Li, Shaogang Gong, Chengjie Wang, Yanwei Fu
HIGHLIGHT: Hence, we start from a generic representation pre-trained by a cross-entropy loss and a conventional distance-based classifier, along with an image retrieval view, to employ a re-ranking process to calibrate a target distance matrix by discovering the k-reciprocal neighbours within the task.

512, TITLE: Robust and Accurate Superquadric Recovery: A Probabilistic Approach
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Robust_and_Accurate_Superquadric_Recovery_A_Probabilistic_Approach_CVPR_2022_paper.html
AUTHORS: Weixiao Liu, Yuwei Wu, Sipu Ruan, Gregory S. Chirikjian
HIGHLIGHT: The superquadric recovery is formulated as a Maximum Likelihood Estimation (MLE) problem. We propose an algorithm, Expectation, Maximization, and Switching (EMS), to solve this problem, where: (1) outliers are predicted from the posterior perspective; (2) the superquadric parameter is optimized by the trust-region reflective algorithm; and (3) local optima are avoided by globally searching and switching among parameters encoding similar superquadrics.

513, TITLE: Zero-Shot Text-Guided Object Generation With Dream Fields
http://openaccess.thecvf.com//content/CVPR2022/html/Jain_Zero-Shot_Text-Guided_Object_Generation_With_Dream_Fields_CVPR_2022_paper.html
AUTHORS: Ajay Jain, Ben Mildenhall, Jonathan T. Barron, Pieter Abbeel, Ben Poole
HIGHLIGHT: To improve fidelity and visual quality, we introduce simple geometric priors, including sparsityinducing transmittance regularization, scene bounds, and new MLP architectures.

514, TITLE: Learning Pixel Trajectories With Multiscale Contrastive Random Walks
http://openaccess.thecvf.com//content/CVPR2022/html/Bian_Learning_Pixel_Trajectories_With_Multiscale_Contrastive_Random_Walks_CVPR_2022_paper.html
AUTHORS: Zhangxing Bian, Allan Jabri, Alexei A. Efros, Andrew Owens
HIGHLIGHT: The main contribution is introducing hierarchy into the search problem by computing the transition matrix in a coarse-to-fine manner, forming a multiscale contrastive random walk.

515, TITLE: Self-Supervised Correlation Mining Network for Person Image Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Self-Supervised_Correlation_Mining_Network_for_Person_Image_Generation_CVPR_2022_paper.html
AUTHORS: Zijian Wang, Xingqun Qi, Kun Yuan, Muyi Sun
HIGHLIGHT: In this paper, we propose a Self-supervised Correlation Mining Network (SCM-Net) to rearrange the source images in the feature space, in which two collaborative modules are integrated, Decomposed Style Encoder (DSE) and Correlation Mining Module (CMM).

516, TITLE: Grounding Answers for Visual Questions Asked by Visually Impaired People
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Grounding_Answers_for_Visual_Questions_Asked_by_Visually_Impaired_People_CVPR_2022_paper.html
AUTHORS: Chongyan Chen, Samreen Anjum, Danna Gurari
HIGHLIGHT: We introduce the VizWiz-VQA-Grounding dataset, the first dataset that visually grounds answers to visual questions asked by people with visual impairments.

517, TITLE: Task Adaptive Parameter Sharing for Multi-Task Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Wallingford_Task_Adaptive_Parameter_Sharing_for_Multi-Task_Learning_CVPR_2022_paper.html
AUTHORS: Matthew Wallingford, Hao Li, Alessandro Achille, Avinash Ravichandran, Charless Fowlkes, Rahul Bhotika, Stefano Soatto

HIGHLIGHT: To efficiently learn multiple downstream tasks we introduce Task Adaptive Parameter Sharing (TAPS), a simple method for tuning a base model to a new task by adaptively modifying a small, task-specific subset of layers.

518, TITLE: Sparse Instance Activation for Real-Time Instance Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Cheng_Sparse_Instance_Activation_for_Real-Time_Instance_Segmentation_CVPR_2022_paper.html
AUTHORS: Tianheng Cheng, Xinggang Wang, Shaoyu Chen, Wenqiang Zhang, Qian Zhang, Chang Huang, Zhaoxiang Zhang, Wenyu Liu
HIGHLIGHT: In this paper, we propose a conceptually novel, efficient, and fully convolutional framework for real-time instance segmentation.

519, TITLE: Automatic Color Image Stitching Using Quaternion Rank-1 Alignment
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Automatic_Color_Image_Stitching_Using_Quaternion_Rank-1_Alignment_CVPR_2022_paper.html
AUTHORS: Jiaxue Li, Yicong Zhou
HIGHLIGHT: To quantitatively evaluate image stitching performance, we propose a perceptual seam quality (PSQ) measure to calculate misalignments of local regions along the seamline.

520, TITLE: VisualGPT: Data-Efficient Adaptation of Pretrained Language Models for Image Captioning
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_VisualGPT_Data-Efficient_Adaptation_of_Pretrained_Language_Models_for_Image_Captioning_CVPR_2022_paper.html
AUTHORS: Jun Chen, Han Guo, Kai Yi, Boyang Li, Mohamed Elhoseiny
HIGHLIGHT: We propose VisualGPT, which employs a novel self-resurrecting encoder-decoder attention mechanism to quickly adapt the PLM with a small amount of in-domain image-text data.

521, TITLE: ESCNet: Gaze Target Detection With the Understanding of 3D Scenes
http://openaccess.thecvf.com//content/CVPR2022/html/Bao_ESCNet_Gaze_Target_Detection_With_the_Understanding_of_3D_Scenes_CVPR_2022_paper.html
AUTHORS: Jun Bao, Buyu Liu, Jun Yu
HIGHLIGHT: This paper aims to address the single image gaze target detection problem.

522, TITLE: Can You Spot the Chameleon? Adversarially Camouflaging Images From Co-Salient Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Gao_Can_You_Spot_the_Chameleon_Adversarially_Camouflaging_Images_From_Co-Salient_CVPR_2022_paper.html
AUTHORS: Ruijun Gao, Qing Guo, Felix Juefei-Xu, Hongkai Yu, Huazhu Fu, Wei Feng, Yang Liu, Song Wang
HIGHLIGHT: However, it inevitably poses an entirely new safety and security issue, i.e., highly personal and sensitive content can potentially be extracting by powerful CoSOD methods. In this paper, we address this problem from the perspective of adversarial attacks and identify a novel task: adversarial co-saliency attack.

523, TITLE: Finding Badly Drawn Bunnies
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Finding_Badly_Drawn_Bunnies_CVPR_2022_paper.html
AUTHORS: Lan Yang, Kaiyue Pang, Honggang Zhang, Yi-Zhe Song
HIGHLIGHT: We propose Geometry-Aware Classification Layer (GACL), a generic method that makes feature-magnitude-as-quality-metric possible and importantly does it without the need for specific quality annotations from humans.

524, TITLE: Point2Cyl: Reverse Engineering 3D Objects From Point Clouds to Extrusion Cylinders
http://openaccess.thecvf.com//content/CVPR2022/html/Uy_Point2Cyl_Reverse_Engineering_3D_Objects_From_Point_Clouds_to_Extrusion_CVPR_2022_paper.html
AUTHORS: Mikaela Angelina Uy, Yen-Yu Chang, Minhyuk Sung, Purvi Goel, Joseph G. Lambourne, Tolga Birdal, Leonidas J. Guibas
HIGHLIGHT: We propose Point2Cyl, a supervised network transforming a raw 3D point cloud to a set of extrusion cylinders.

525, TITLE: All-Photon Polarimetric Time-of-Flight Imaging
http://openaccess.thecvf.com//content/CVPR2022/html/Baek_All-Photon_Polarimetric_Time-of-Flight_Imaging_CVPR_2022_paper.html
AUTHORS: Seung-Hwan Baek, Felix Heide
HIGHLIGHT: We depart from the principle of first arrival and instead propose an all-photon ToF imaging method relying on polarization changes that analyzes both first- and late-arriving photons for shape and material scene understanding.

526, TITLE: MHFormer: Multi-Hypothesis Transformer for 3D Human Pose Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Li_MHFormer_Multi-Hypothesis_Transformer_for_3D_Human_Pose_Estimation_CVPR_2022_paper.html
AUTHORS: Wenhao Li, Hong Liu, Hao Tang, Pichao Wang, Luc Van Gool
HIGHLIGHT: However, those works ignore the fact that it is an inverse problem where multiple feasible solutions (i.e., hypotheses) exist. To relieve this limitation, we propose a Multi-Hypothesis Transformer (MHFormer) that learns spatio-temporal representations of multiple plausible pose hypotheses.

527, TITLE: Surface-Aligned Neural Radiance Fields for Controllable 3D Human Synthesis
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_Surface-Aligned_Neural_Radiance_Fields_for_Controllable_3D_Human_Synthesis_CVPR_2022_paper.html

AUTHORS: Tianhan Xu, Yasuhiro Fujita, Eiichi Matsumoto
HIGHLIGHT: We propose a new method for reconstructing controllable implicit 3D human models from sparse multi-view RGB videos.

528, TITLE: Learning From Temporal Gradient for Semi-Supervised Action Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Xiao_Learning_From_Temporal_Gradient_for_Semi-Supervised_Action_Recognition_CVPR_2022_paper.html
AUTHORS: Junfei Xiao, Longlong Jing, Lin Zhang, Ju He, Qi She, Zongwei Zhou, Alan Yuille, Yingwei Li
HIGHLIGHT: To better leverage the encoded temporal information in videos, we introduce temporal gradient as an additional modality for more attentive feature extraction in this paper.

529, TITLE: Towards Implicit Text-Guided 3D Shape Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Towards_Implicit_Text-Guided_3D_Shape_Generation_CVPR_2022_paper.html
AUTHORS: Zhengzhe Liu, Yi Wang, Xiaojuan Qi, Chi-Wing Fu
HIGHLIGHT: In this work, we explore the challenging task of generating 3D shapes from text.

530, TITLE: Audio-Driven Neural Gesture Reenactment With Video Motion Graphs
http://openaccess.thecvf.com//content/CVPR2022/html/Zhou_Audio-Driven_Neural_Gesture_Reenactment_With_Video_Motion_Graphs_CVPR_2022_paper.html
AUTHORS: Yang Zhou, Jimei Yang, Dingzeyu Li, Jun Saito, Deepali Aneja, Evangelos Kalogerakis
HIGHLIGHT: We present a method that reenacts a high-quality video with gestures matching a target speech audio.

531, TITLE: SoftCollage: A Differentiable Probabilistic Tree Generator for Image Collage
http://openaccess.thecvf.com//content/CVPR2022/html/Yu_SoftCollage_A_Differentiable_Probabilistic_Tree_Generator_for_Image_Collage_CVPR_2022_paper.html
AUTHORS: Jiahao Yu, Li Chen, Mingrui Zhang, Mading Li
HIGHLIGHT: We propose SoftCollage, a novel method that employs a neural-based differentiable probabilistic tree generator to produce the probability distribution of correlation-preserving collage tree conditioned on deep image feature, aspect ratio and canvas size.

532, TITLE: Transforming Model Prediction for Tracking
http://openaccess.thecvf.com//content/CVPR2022/html/Mayer_Transforming_Model_Prediction_for_Tracking_CVPR_2022_paper.html
AUTHORS: Christoph Mayer, Martin Danelljan, Goutam Bhat, Matthieu Paul, Danda Pani Paudel, Fisher Yu, Luc Van Gool
HIGHLIGHT: While this inductive bias integrates valuable domain knowledge, it limits the expressivity of the tracking network. In this work, we therefore propose a tracker architecture employing a Transformer-based model prediction module.

533, TITLE: A Unified Framework for Implicit Sinkhorn Differentiation
http://openaccess.thecvf.com//content/CVPR2022/html/Eisenberger_A_Unified_Framework_for_Implicit_Sinkhorn_Differentiation_CVPR_2022_paper.html
AUTHORS: Marvin Eisenberger, Aysim Toker, Laura Leal-Taixé, Florian Bernard, Daniel Cremers
HIGHLIGHT: To allow for an efficient training of respective neural networks, we propose an algorithm that obtains analytical gradients of a Sinkhorn layer via implicit differentiation.

534, TITLE: DGECN: A Depth-Guided Edge Convolutional Network for End-to-End 6D Pose Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Cao_DGECN_A_Depth-Guided_Edge_Convolutional_Network_for_End-to-End_6D_Pose_CVPR_2022_paper.html
AUTHORS: Tuo Cao, Fei Luo, Yanping Fu, Wenxiao Zhang, Shengjie Zheng, Chunxia Xiao
HIGHLIGHT: However, most of them hardly consider the geometric features in 3D space, and ignore the topology cues when performing differentiable RANSAC algorithms. To this end, we proposed a Depth-Guided Edge Convolutional Network (DGECN) for 6D pose estimation task.

535, TITLE: Unsupervised Vision-Language Parsing: Seamlessly Bridging Visual Scene Graphs With Language Structures via Dependency Relationships
http://openaccess.thecvf.com//content/CVPR2022/html/Lou_Unsupervised_Vision-Language_Parsing_Seamlessly_Bridging_Visual_Scene_Graphs_With_Language_CVPR_2022_paper.html
AUTHORS: Chao Lou, Wenjuan Han, Yuhuan Lin, Zilong Zheng
HIGHLIGHT: However, how to construct a joint vision-language (VL) structure has barely been investigated. More challenging but worthwhile, we introduce a new task that targets on inducing such a joint VL structure in an unsupervised manner.

536, TITLE: Open-Vocabulary Instance Segmentation via Robust Cross-Modal Pseudo-Labeling
http://openaccess.thecvf.com//content/CVPR2022/html/Huynh_Open-Vocabulary_Instance_Segmentation_via_Robust_Cross-Modal_Pseudo-Labeling_CVPR_2022_paper.html
AUTHORS: Dat Huynh, Jason Kuen, Zhe Lin, Jiuxiang Gu, Ehsan Elhamifar
HIGHLIGHT: However, the high-level textual information learned from caption pretraining alone cannot effectively encode the details required for pixel-wise segmentation. To address this, we propose a cross-modal pseudo-labeling framework, which generates training pseudo masks by aligning word semantics in captions with visual features of object masks in images.

537, TITLE: Locality-Aware Inter- and Intra-Video Reconstruction for Self-Supervised Correspondence Learning

http://openaccess.thecvf.com//content/CVPR2022/html/Li_Locality-Aware_Inter-_and_Intra-Video_Reconstruction_for_Self-Supervised_Correspondence_Learning_CVPR_2022_paper.html
AUTHORS: Liulei Li, Tianfei Zhou, Wenguan Wang, Lu Yang, Jianwu Li, Yi Yang
HIGHLIGHT: We develop LIIR, a locality-aware inter-and intra-video reconstruction framework that fills in three missing pieces, i.e., instance discrimination, location awareness, and spatial compactness, of self-supervised correspondence learning puzzle.

538, TITLE: A Versatile Multi-View Framework for LiDAR-Based 3D Object Detection With Guidance From Panoptic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Fazlali_A_Versatile_Multi-View_Framework_for_LiDAR-Based_3D_Object_Detection_With_CVPR_2022_paper.html
AUTHORS: Hamidreza Fazlali, Yixuan Xu, Yuan Ren, Bingbing Liu
HIGHLIGHT: In this paper, we propose a novel multi-task framework that jointly performs 3D object detection and panoptic segmentation.

539, TITLE: Query and Attention Augmentation for Knowledge-Based Explainable Reasoning
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Query_and_Attention_Augmentation_for_Knowledge-Based_Explainable_Reasoning_CVPR_2022_paper.html
AUTHORS: Yifeng Zhang, Ming Jiang, Qi Zhao
HIGHLIGHT: To bridge this research gap, we present Query and Attention Augmentation, a general approach that augments neural module networks to jointly reason about visual and external knowledge.

540, TITLE: Winoground: Probing Vision and Language Models for Visio-Linguistic Compositionality
http://openaccess.thecvf.com//content/CVPR2022/html/Thrush_Winoground_Probing_Vision_and_Language_Models_for_Visio-Linguistic_Compositionality_CVPR_2022_paper.html
AUTHORS: Tristan Thrush, Ryan Jiang, Max Bartolo, Amanpreet Singh, Adina Williams, Douwe Kiela, Candace Ross
HIGHLIGHT: We present a novel task and dataset for evaluating the ability of vision and language models to conduct visio-linguistic compositional reasoning, which we call Winoground.

541, TITLE: RFNet: Unsupervised Network for Mutually Reinforcing Multi-Modal Image Registration and Fusion
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_RFNet_Unsupervised_Network_for_Mutually_Reinforcing_Multi-Modal_Image_Registration_and_CVPR_2022_paper.html
AUTHORS: Han Xu, Jiayi Ma, Jiteng Yuan, Zhuliang Le, Wei Liu
HIGHLIGHT: In this paper, we propose a novel method to realize multi-modal image registration and fusion in a mutually reinforcing framework, termed as RFNet.

542, TITLE: Progressive Attention on Multi-Level Dense Difference Maps for Generic Event Boundary Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Tang_Progressive_Attention_on_Multi-Level_Dense_Difference_Maps_for_Generic_Event_CVPR_2022_paper.html
AUTHORS: Jiaqi Tang, Zhaoyang Liu, Chen Qian, Wayne Wu, Limin Wang
HIGHLIGHT: The main challenge of this task is perceiving various temporal variations of diverse event boundaries. To this end, this paper presents an effective and end-to-end learnable framework (DDM-Net).

543, TITLE: Interactron: Embodied Adaptive Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Kotar_Interactron_Embodied_Adaptive_Object_Detection_CVPR_2022_paper.html
AUTHORS: Klemen Kotar, Roozbeh Mottaghi
HIGHLIGHT: In this paper, we propose Interactron, a method for adaptive object detection in an interactive setting, where the goal is to perform object detection in images observed by an embodied agent navigating in different environments.

544, TITLE: 3D Scene Painting via Semantic Image Synthesis
http://openaccess.thecvf.com//content/CVPR2022/html/Jeong_3D_Scene_Painting_via_Semantic_Image_Synthesis_CVPR_2022_paper.html
AUTHORS: Jaebong Jeong, Janghun Jo, Sunghyun Cho, Jaesik Park
HIGHLIGHT: We propose a novel approach to 3D scene painting using a configurable 3D scene layout.

545, TITLE: MeMOT: Multi-Object Tracking With Memory
http://openaccess.thecvf.com//content/CVPR2022/html/Cai_MeMOT_Multi-Object_Tracking_With_Memory_CVPR_2022_paper.html
AUTHORS: Jiarui Cai, Mingze Xu, Wei Li, Yuanjun Xiong, Wei Xia, Zhuowen Tu, Stefano Soatto
HIGHLIGHT: We propose an online tracking algorithm that performs the object detection and data association under a common framework, capable of linking objects after a long time span.

546, TITLE: Revisiting Weakly Supervised Pre-Training of Visual Perception Models
http://openaccess.thecvf.com//content/CVPR2022/html/Singh_Revisiting_Weakly_Supervised_Pre-Training_of_Visual_Perception_Models_CVPR_2022_paper.html
AUTHORS: Mannat Singh, Laura Gustafson, Aaron Adcock, Vinicius de Freitas Reis, Bugra Gedik, Raj Prateek Kosaraju, Dhruv Mahajan, Ross Girshick, Piotr Dollár, Laurens van der Maaten
HIGHLIGHT: This paper revisits weakly-supervised pre-training of models using hashtag supervision with modern versions of residual networks and the largest-ever dataset of images and corresponding hashtags.

547, TITLE: Semi-Supervised Semantic Segmentation With Error Localization Network
http://openaccess.thecvf.com//content/CVPR2022/html/Kwon_Semi-Supervised_Semantic_Segmentation_With_Error_Localization_Network_CVPR_2022_paper.html
AUTHORS: Donghyeon Kwon, Suha Kwak
HIGHLIGHT: The unlabeled images are usually assigned pseudo labels to be used in training, which however often causes the risk of performance degradation due to the confirmation bias towards errors on the pseudo labels. We present a novel method that resolves this chronic issue of pseudo labeling.

548, TITLE: Meta Convolutional Neural Networks for Single Domain Generalization
http://openaccess.thecvf.com//content/CVPR2022/html/Wan_Meta_Convolutional_Neural_Networks_for_Single_Domain_Generalization_CVPR_2022_paper.html
AUTHORS: Chaoqun Wan, Xu Shen, Yonggang Zhang, Zhiheng Yin, Xinmei Tian, Feng Gao, Jianqiang Huang, Xian-Sheng Hua
HIGHLIGHT: In this paper, we propose a new model, termed meta convolutional neural network, to solve the single domain generalization problem in image recognition.

549, TITLE: Generalizing Gaze Estimation With Rotation Consistency
http://openaccess.thecvf.com//content/CVPR2022/html/Bao_Generalizing_Gaze_Estimation_With_Rotation_Consistency_CVPR_2022_paper.html
AUTHORS: Yiwei Bao, Yunfei Liu, Haofei Wang, Feng Lu
HIGHLIGHT: In this paper, we discover the rotation-consistency property in gaze estimation and introduce the 'sub-label' for unsupervised domain adaptation.

550, TITLE: Anomaly Detection via Reverse Distillation From One-Class Embedding
http://openaccess.thecvf.com//content/CVPR2022/html/Deng_Anomaly_Detection_via_Reverse_Distillation_From_One-Class_Embedding_CVPR_2022_paper.html
AUTHORS: Hanqiu Deng, Xingyu Li
HIGHLIGHT: However, using similar or identical architectures to build the teacher and student models in previous studies hinders the diversity of anomalous representations. To tackle this problem, we propose a novel T-S model consisting of a teacher encoder and a student decoder and introduce a simple yet effective "reverse distillation" paradigm accordingly.

551, TITLE: Fine-Grained Object Classification via Self-Supervised Pose Alignment
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Fine-Grained_Object_Classification_via_Self-Supervised_Pose_Alignment_CVPR_2022_paper.html
AUTHORS: Xuhui Yang, Yaowei Wang, Ke Chen, Yong Xu, Yonghong Tian
HIGHLIGHT: For discounting pose variations, this paper proposes to learn a novel graph based object representation to reveal a global configuration of local parts for self-supervised pose alignment across classes, which is employed as an auxiliary feature regularization on a deep representation learning network.

552, TITLE: Spatio-Temporal Gating-Adjacency GCN for Human Motion Prediction
http://openaccess.thecvf.com//content/CVPR2022/html/Zhong_Spatio-Temporal_Gating-Adjacency_GCN_for_Human_Motion_Prediction_CVPR_2022_paper.html
AUTHORS: Chongyang Zhong, Lei Hu, Zihao Zhang, Yongjing Ye, Shihong Xia
HIGHLIGHT: However, considering the variants and diverse action types in human motion data, the cross-dependency of the spatio-temporal relationships will be difficult to depict due to the decoupled modeling strategy, which may also exacerbate the problem of insufficient generalization. Therefore, we propose the Spatio-Temporal Gating-Adjacency GCN(GAGCN) to learn the complex spatio-temporal dependencies over diverse action types.

553, TITLE: CellTypeGraph: A New Geometric Computer Vision Benchmark
http://openaccess.thecvf.com//content/CVPR2022/html/Cerrone_CellTypeGraph_A_New_Geometric_Computer_Vision_Benchmark_CVPR_2022_paper.html
AUTHORS: Lorenzo Cerrone, Athul Vijayan, Tejasvinee Mody, Kay Schneitz, Fred A. Hamprecht
HIGHLIGHT: Classifying all cells in an organ is a relevant and difficult problem from plant developmental biology. We here abstract the problem into a new benchmark for node classification in a geo-referenced graph.

554, TITLE: Clustering Plotted Data by Image Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Naous_Clustering_Plotted_Data_by_Image_Segmentation_CVPR_2022_paper.html
AUTHORS: Tarek Naous, Srinjay Sarkar, Abubakar Abid, James Zou
HIGHLIGHT: In this paper, we present a different way of clustering points in 2-dimensional space, inspired by how humans cluster data: by training neural networks to perform instance segmentation on plotted data.

555, TITLE: Accelerating Neural Network Optimization Through an Automated Control Theory Lens
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Accelerating_Neural_Network_Optimization_Through_an_Automated_Control_Theory_Lens_CVPR_2022_paper.html
AUTHORS: Jiahao Wang, Baoyuan Wu, Rui Su, Mingdeng Cao, Shuwei Shi, Wanli Ouyang, Yujiu Yang
HIGHLIGHT: This paper studies the optimizer for accelerating the time-consuming deep network training through an automated control theory lens.

556, TITLE: Animal Kingdom: A Large and Diverse Dataset for Animal Behavior Understanding

http://openaccess.thecvf.com//content/CVPR2022/html/Ng_Animal_Kingdom_A_Large_and_Diverse_Dataset_for_Animal_Behavior_CVPR_2022_paper.html
AUTHORS: Xun Long Ng, Kian Eng Ong, Qichen Zheng, Yun Ni, Si Yong Yeo, Jun Liu
HIGHLIGHT: However, existing animal behavior datasets have limitations in multiple aspects, including limited numbers of animal classes, data samples and provided tasks, and also limited variations in environmental conditions and viewpoints. To address these limitations, we create a large and diverse dataset, Animal Kingdom, that provides multiple annotated tasks to enable a more thorough understanding of natural animal behaviors.

557, TITLE: Learning To Learn Across Diverse Data Biases in Deep Face Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Learning_To_Learn_Across_Diverse_Data_Biases_in_Deep_Face_CVPR_2022_paper.html
AUTHORS: Chang Liu, Xiang Yu, Yi-Hsuan Tsai, Masoud Faraki, Ramin Moslemi, Manmohan Chandraker, Yun Fu
HIGHLIGHT: In this paper, we show that many bias variations such as ethnicity, head pose, occlusion and blur can jointly affect the accuracy significantly.

558, TITLE: Back to Reality: Weakly-Supervised 3D Object Detection With Shape-Guided Label Enhancement
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_Back_to_Reality_Weakly-Supervised_3D_Object_Detection_With_Shape-Guided_Label_CVPR_2022_paper.html
AUTHORS: Xiuwei Xu, Yifan Wang, Yu Zheng, Yongming Rao, Jie Zhou, Jiwen Lu
HIGHLIGHT: In this paper, we propose a weakly-supervised approach for 3D object detection, which makes it possible to train a strong 3D detector with position-level annotations (i.e. annotations of object centers).

559, TITLE: Long-Tail Recognition via Compositional Knowledge Transfer
http://openaccess.thecvf.com//content/CVPR2022/html/Parisot_Long-Tail_Recognition_via_Compositional_Knowledge_Transfer_CVPR_2022_paper.html
AUTHORS: Sarah Parisot, Pedro M. Esperança, Steven McDonagh, Tamas J. Madarasz, Yongxin Yang, Zhenguo Li
HIGHLIGHT: In this work, we introduce a novel strategy for long-tail recognition that addresses the tail classes' few-shot problem via training-free knowledge transfer.

560, TITLE: EI-CLIP: Entity-Aware Interventional Contrastive Learning for E-Commerce Cross-Modal Retrieval
http://openaccess.thecvf.com//content/CVPR2022/html/Ma_EI-CLIP_Entity-Aware_Interventional_Contrastive_Learning_for_E-Commerce_Cross-Modal_Retrieval_CVPR_2022_paper.html
AUTHORS: Haoyu Ma, Handong Zhao, Zhe Lin, Ajinkya Kale, Zhangyang Wang, Tong Yu, Jiuxiang Gu, Sunav Choudhary, Xiaohui Xie
HIGHLIGHT: By formulating such out-of-distribution finetuning process in the Causal Inference paradigm, we view the erroneous semantics of these special entities as confounders to cause the retrieval failure. To rectify these semantics for aligning with e-commerce domain knowledge, we propose an intervention-based entity-aware contrastive learning framework with two modules, i.e., the Confounding Entity Selection Module and Entity-Aware Learning Module.

561, TITLE: Multi-Dimensional, Nuanced and Subjective - Measuring the Perception of Facial Expressions
http://openaccess.thecvf.com//content/CVPR2022/html/Bryant_Multi-Dimensional_Nuanced_and_Subjective_-_Measuring_the_Perception_of_Facial_CVPR_2022_paper.html
AUTHORS: De'Aira Bryant, Siqi Deng, Nashlie Sephus, Wei Xia, Pietro Perona
HIGHLIGHT: We propose a methodology for collecting and modeling multidimensional modulated expression annotations from human annotators.

562, TITLE: PyMiceTracking: An Open-Source Toolbox for Real-Time Behavioral Neuroscience Experiments
http://openaccess.thecvf.com//content/CVPR2022/html/Menezes_PyMiceTracking_An_Open-Source_Toolbox_for_Real-Time_Behavioral_Neuroscience_Experiments_CVPR_2022_paper.html
AUTHORS: Richardson Menezes, Aron de Miranda, Helton Maia
HIGHLIGHT: PyMiceTracking: An Open-Source Toolbox for Real-Time Behavioral Neuroscience Experiments

563, TITLE: Self-Taught Metric Learning Without Labels
http://openaccess.thecvf.com//content/CVPR2022/html/Kim_Self-Taught_Metric_Learning_Without_Labels_CVPR_2022_paper.html
AUTHORS: Sungyeon Kim, Dongwon Kim, Minsu Cho, Suha Kwak
HIGHLIGHT: We present a novel self-taught framework for unsupervised metric learning, which alternates between predicting class-equivalence relations between data through a moving average of an embedding model and learning the model with the predicted relations as pseudo labels.

564, TITLE: MeMViT: Memory-Augmented Multiscale Vision Transformer for Efficient Long-Term Video Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Wu_MeMViT_Memory-Augmented_Multiscale_Vision_Transformer_for_Efficient_Long-Term_Video_Recognition_CVPR_2022_paper.html
AUTHORS: Chao-Yuan Wu, Yanghao Li, Karttikeya Mangalam, Haoqi Fan, Bo Xiong, Jitendra Malik, Christoph Feichtenhofer
HIGHLIGHT: Most existing video architectures can only process <5 seconds of a video without hitting the computation or memory bottlenecks. In this paper, we propose a new strategy to overcome this challenge.

565, TITLE: Fine-Grained Temporal Contrastive Learning for Weakly-Supervised Temporal Action Localization

http://openaccess.thecvf.com//content/CVPR2022/html/Gao_Fine-Grained_Temporal_Contrastive_Learning_for_Weakly-Supervised_Temporal_Action_Localization_CVPR_2022_paper.html
AUTHORS: Junyu Gao, Mengyuan Chen, Changsheng Xu
HIGHLIGHT: This paper argues that learning by contextually comparing sequence-to-sequence distinctions offers an essential inductive bias in WSAL and helps identify coherent action instances.

566, TITLE: Embracing Single Stride 3D Object Detector With Sparse Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Fan_Embracing_Single_Stride_3D_Object_Detector_With_Sparse_Transformer_CVPR_2022_paper.html
AUTHORS: Lue Fan, Ziqi Pang, Tianyuan Zhang, Yu-Xiong Wang, Hang Zhao, Feng Wang, Naiyan Wang, Zhaoxiang Zhang
HIGHLIGHT: Overlooking this difference, many 3D detectors directly follow the common practice of 2D detectors, which downsample the feature maps even after quantizing the point clouds. In this paper, we start by rethinking how such multi-stride stereotype affects the LiDAR-based 3D object detectors.

567, TITLE: Multidimensional Belief Quantification for Label-Efficient Meta-Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Pandey_Multidimensional_Belief_Quantification_for_Label-Efficient_Meta-Learning_CVPR_2022_paper.html
AUTHORS: Deep Shankar Pandey, Qi Yu
HIGHLIGHT: We propose a novel uncertainty-aware task selection model for label efficient meta-learning.

568, TITLE: UTC: A Unified Transformer With Inter-Task Contrastive Learning for Visual Dialog
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_UTC_A_Unified_Transformer_With_Inter-Task_Contrastive_Learning_for_Visual_CVPR_2022_paper.html
AUTHORS: Cheng Chen, Zhenshan Tan, Qingrong Cheng, Xin Jiang, Qun Liu, Yudong Zhu, Xiaodong Gu
HIGHLIGHT: In this paper, we propose a contrastive learning-based framework UTC to unify and facilitate both discriminative and generative tasks in visual dialog with a single model.

569, TITLE: Relieving Long-Tailed Instance Segmentation via Pairwise Class Balance
http://openaccess.thecvf.com//content/CVPR2022/html/He_Relieving_Long-Tailed_Instance_Segmentation_via_Pairwise_Class_Balance_CVPR_2022_paper.html
AUTHORS: Yin-Yin He, Peizhen Zhang, Xiu-Shen Wei, Xiangyu Zhang, Jian Sun
HIGHLIGHT: In this paper, we explore to excavate the confusion matrix, which carries the fine-grained misclassification details, to relieve the pairwise biases, generalizing the coarse one.

570, TITLE: Online Convolutional Re-Parameterization
http://openaccess.thecvf.com//content/CVPR2022/html/Hu_Online_Convolutional_Re-Parameterization_CVPR_2022_paper.html
AUTHORS: Mu Hu, Junyi Feng, Jiashen Hua, Baisheng Lai, Jianqiang Huang, Xiaojin Gong, Xian-Sheng Hua
HIGHLIGHT: In this paper, we present online convolutional re-parameterization (OREPA), a two-stage pipeline, aiming to reduce the huge training overhead by squeezing the complex training-time block into a single convolution.

571, TITLE: Mimicking the Oracle: An Initial Phase Decorrelation Approach for Class Incremental Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Shi_Mimicking_the_Oracle_An_Initial_Phase_Decorrelation_Approach_for_Class_CVPR_2022_paper.html
AUTHORS: Yujun Shi, Kuangqi Zhou, Jian Liang, Zihang Jiang, Jiashi Feng, Philip H.S. Torr, Song Bai, Vincent Y. F. Tan
HIGHLIGHT: We find that, with fewer training classes, the data representations of each class lie in a long and narrow region; with more training classes, the representations of each class scatter more uniformly. Inspired by this observation, we propose Class-wise Decorrelation (CwD) that effectively regularizes representations of each class to scatter more uniformly, thus mimicking the model jointly trained with all classes (i.e., the oracle model).

572, TITLE: RIDDLE: Lidar Data Compression With Range Image Deep Delta Encoding
http://openaccess.thecvf.com//content/CVPR2022/html/Zhou_RIDDLE_Lidar_Data_Compression_With_Range_Image_Deep_Delta_Encoding_CVPR_2022_paper.html
AUTHORS: Xuanyu Zhou, Charles R. Qi, Yin Zhou, Dragomir Anguelov
HIGHLIGHT: In this work, we show that directly compressing the range images can leverage the lidar scanning pattern, compared to compressing the unprojected point clouds.

573, TITLE: RelTransformer: A Transformer-Based Long-Tail Visual Relationship Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_RelTransformer_A_Transformer-Based_Long-Tail_Visual_Relationship_Recognition_CVPR_2022_paper.html
AUTHORS: Jun Chen, Aniket Agarwal, Sherif Abdelkarim, Deyao Zhu, Mohamed Elhoseiny
HIGHLIGHT: This paper shows that modeling an effective message-passing flow through an attention mechanism can be critical to tackling the compositionality and long-tail challenges in VRR.

574, TITLE: HODEC: Towards Efficient High-Order DEcomposed Convolutional Neural Networks
http://openaccess.thecvf.com//content/CVPR2022/html/Yin_HODEC_Towards_Efficient_High-Order_DEcomposed_Convolutional_Neural_Networks_CVPR_2022_paper.html
AUTHORS: Miao Yin, Yang Sui, Wanzhao Yang, Xiao Zang, Yu Gong, Bo Yuan
HIGHLIGHT: However, many of the existing solutions, though can efficiently reduce CNN model sizes, are very difficult to bring considerable saving for computational costs, especially when the compression ratio is not huge, thereby causing the severe

computation inefficiency problem. To overcome this challenge, in this paper we propose efficient High-Order DEcomposed Convolution (HODEC).

575, TITLE: RigidFlow: Self-Supervised Scene Flow Learning on Point Clouds by Local Rigidity Prior
http://openaccess.thecvf.com//content/CVPR2022/html/Li_RigidFlow_Self-Supervised_Scene_Flow_Learning_on_Point_Clouds_by_Local_CVPR_2022_paper.html
AUTHORS: Ruibo Li, Chi Zhang, Guosheng Lin, Zhe Wang, Chunhua Shen
HIGHLIGHT: In this work, we focus on scene flow learning on point clouds in a self-supervised manner.

576, TITLE: Smooth Maximum Unit: Smooth Activation Function for Deep Networks Using Smoothing Maximum Technique
http://openaccess.thecvf.com//content/CVPR2022/html/Biswas_Smooth_Maximum_Unit_Smooth_Activation_Function_for_Deep_Networks_Using_CVPR_2022_paper.html
AUTHORS: Koushik Biswas, Sandeep Kumar, Shilpak Banerjee, Ashish Kumar Pandey
HIGHLIGHT: Rectified Linear Unit (ReLU) is a popular hand-designed activation function and is the most common choice in the deep learning community due to its simplicity though ReLU has some drawbacks. In this paper, we have proposed two new novel activation functions based on approximation of the maximum function, and we call these functions Smooth Maximum Unit (SMU and SMU-1).

577, TITLE: Learning Invisible Markers for Hidden Codes in Offline-to-Online Photography
http://openaccess.thecvf.com//content/CVPR2022/html/Jia_Learning_Invisible_Markers_for_Hidden_Codes_in_Offline-to-Online_Photography_CVPR_2022_paper.html
AUTHORS: Jun Jia, Zhongpai Gao, Dandan Zhu, Xiongkuo Min, Guangtao Zhai, Xiaokang Yang
HIGHLIGHT: This paper proposes a novel invisible information hiding architecture for display/print-camera scenarios, consisting of hiding, locating, correcting, and recovery, where invisible markers are learned to make hidden codes truly invisible.

578, TITLE: Personalized Image Aesthetics Assessment With Rich Attributes
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Personalized_Image_Aesthetics_Assessment_With_Rich_Attributes_CVPR_2022_paper.html
AUTHORS: Yuzhe Yang, Liwu Xu, Leida Li, Nan Qie, Yaqian Li, Peng Zhang, Yandong Guo
HIGHLIGHT: To solve the dilemma, we conduct so far, the most comprehensive subjective study of personalized image aesthetics and introduce a new Personalized image Aesthetics database with Rich Attributes (PARA), which consists of 31,220 images with annotations by 438 subjects.

579, TITLE: Task2Sim: Towards Effective Pre-Training and Transfer From Synthetic Data
http://openaccess.thecvf.com//content/CVPR2022/html/Mishra_Task2Sim_Towards_Effective_Pre-Training_and_Transfer_From_Synthetic_Data_CVPR_2022_paper.html
AUTHORS: Samarth Mishra, Rameswar Panda, Cheng Perng Phoo, Chun-Fu (Richard) Chen, Leonid Karlinsky, Kate Saenko, Venkatesh Saligrama, Rogerio S. Feris
HIGHLIGHT: It is thus better to tailor synthetic pre-training data to a specific downstream task, for best performance. We introduce Task2Sim, a unified model mapping downstream task representations to optimal simulation parameters to generate synthetic pre-training data for them.

580, TITLE: Part-Based Pseudo Label Refinement for Unsupervised Person Re-Identification
http://openaccess.thecvf.com//content/CVPR2022/html/Cho_Part-Based_Pseudo_Label_Refinement_for_Unsupervised_Person_Re-Identification_CVPR_2022_paper.html
AUTHORS: Yoonki Cho, Woo Jae Kim, Seunghoon Hong, Sung-Eui Yoon
HIGHLIGHT: In this paper, we propose a novel Part-based Pseudo Label Refinement (PPLR) framework that reduces the label noise by employing the complementary relationship between global and part features.

581, TITLE: Bridging the Gap Between Learning in Discrete and Continuous Environments for Vision-and-Language Navigation
http://openaccess.thecvf.com//content/CVPR2022/html/Hong_Bridging_the_Gap_Between_Learning_in_Discrete_and_Continuous_Environments_CVPR_2022_paper.html
AUTHORS: Yicong Hong, Zun Wang, Qi Wu, Stephen Gould
HIGHLIGHT: To bridge the discrete-to-continuous gap, we propose a predictor to generate a set of candidate waypoints during navigation, so that agents designed with high-level actions can be transferred to and trained in continuous environments.

582, TITLE: HDNet: High-Resolution Dual-Domain Learning for Spectral Compressive Imaging
http://openaccess.thecvf.com//content/CVPR2022/html/Hu_HDNet_High-Resolution_Dual-Domain_Learning_for_Spectral_Compressive_Imaging_CVPR_2022_paper.html
AUTHORS: Xiaowan Hu, Yuanhao Cai, Jing Lin, Haoqian Wang, Xin Yuan, Yulun Zhang, Radu Timofte, Luc Van Gool
HIGHLIGHT: So we propose a high-resolution dual-domain learning network (HDNet) for HSI reconstruction.

583, TITLE: OW-DETR: Open-World Detection Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Gupta_OW-DETR_Open-World_Detection_Transformer_CVPR_2022_paper.html
AUTHORS: Akshita Gupta, Sanath Narayan, K J Joseph, Salman Khan, Fahad Shahbaz Khan, Mubarak Shah
HIGHLIGHT: Here, we introduce a novel end-to-end transformer-based framework, OW-DETR, for open-world object detection.

584, TITLE: Learning Deep Implicit Functions for 3D Shapes With Dynamic Code Clouds
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Learning_Deep_Implicit_Functions_for_3D_Shapes_With_Dynamic_Code_CVPR_2022_paper.html
AUTHORS: Tianyang Li, Xin Wen, Yu-Shen Liu, Hua Su, Zhizhong Han
HIGHLIGHT: However, the local codes are constrained at discrete and regular positions like grid points, which makes the code positions difficult to be optimized and limits their representation ability. To solve this problem, we propose to learn DIF with Dynamic Code Cloud, named DCC-DIF.

585, TITLE: Reversible Vision Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Mangalam_Reversible_Vision_Transformers_CVPR_2022_paper.html
AUTHORS: Karttikeya Mangalam, Haoqi Fan, Yanghao Li, Chao-Yuan Wu, Bo Xiong, Christoph Feichtenhofer, Jitendra Malik
HIGHLIGHT: We present Reversible Vision Transformers, a memory efficient architecture design for visual recognition.

586, TITLE: Amodal Panoptic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Mohan_Amodal_Panoptic_Segmentation_CVPR_2022_paper.html
AUTHORS: Rohit Mohan, Abhinav Valada
HIGHLIGHT: This ability of amodal perception forms the basis of our perceptual and cognitive understanding of our world. To enable robots to reason with this capability, we formulate and propose a novel task that we name amodal panoptic segmentation.

587, TITLE: Gravitationally Lensed Black Hole Emission Tomography
http://openaccess.thecvf.com//content/CVPR2022/html/Levis_Gravitationally_Lensed_Black_Hole_Emission_Tomography_CVPR_2022_paper.html
AUTHORS: Aviad Levis, Pratul P. Srinivasan, Andrew A. Chael, Ren Ng, Katherine L. Bouman
HIGHLIGHT: In this work, we propose BH-NeRF, a novel tomography approach that leverages gravitational lensing to recover the continuous 3D emission field near a black hole.

588, TITLE: 3D-Aware Image Synthesis via Learning Structural and Textural Representations
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_3D-Aware_Image_Synthesis_via_Learning_Structural_and_Textural_Representations_CVPR_2022_paper.html
AUTHORS: Yinghao Xu, Sida Peng, Ceyuan Yang, Yujun Shen, Bolei Zhou
HIGHLIGHT: Meanwhile, NeRF is built on volume rendering which can be too costly to produce high-resolution results, increasing the optimization difficulty. To alleviate these two problems, we propose a novel framework, termed as VolumeGAN, for high-fidelity 3D-aware image synthesis, through explicitly learning a structural representation and a textural representation.

589, TITLE: Text-to-Image Synthesis Based on Object-Guided Joint-Decoding Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Wu_Text-to-Image_Synthesis_Based_on_Object-Guided_Joint-Decoding_Transformer_CVPR_2022_paper.html
AUTHORS: Fuxiang Wu, Liu Liu, Fusheng Hao, Fengxiang He, Jun Cheng
HIGHLIGHT: In this paper, we propose an object-guided joint-decoding module to simultaneously generate the image and the corresponding layout.

590, TITLE: Correlation Verification for Image Retrieval
http://openaccess.thecvf.com//content/CVPR2022/html/Lee_Correlation_Verification_for_Image_Retrieval_CVPR_2022_paper.html
AUTHORS: Seongwon Lee, Hongje Seong, Suhyeon Lee, Euntai Kim
HIGHLIGHT: In this study, we propose a novel image retrieval re-ranking network named Correlation Verification Networks (CVNet).

591, TITLE: Unsupervised Vision-and-Language Pre-Training via Retrieval-Based Multi-Granular Alignment
http://openaccess.thecvf.com//content/CVPR2022/html/Zhou_Unsupervised_Vision-and-Language_Pre-Training_via_Retrieval-Based_Multi-Granular_Alignment_CVPR_2022_paper.html
AUTHORS: Mingyang Zhou, Licheng Yu, Amanpreet Singh, Mengjiao Wang, Zhou Yu, Ning Zhang
HIGHLIGHT: In this paper, we propose unsupervised Vision-and-Language pre-training (UVLP) to learn the cross-modal representation from non-parallel image and text datasets.

592, TITLE: Protecting Facial Privacy: Generating Adversarial Identity Masks via Style-Robust Makeup Transfer
http://openaccess.thecvf.com//content/CVPR2022/html/Hu_Protecting_Facial_Privacy_Generating_Adversarial_Identity_Masks_via_Style-Robust_Makeup_CVPR_2022_paper.html
AUTHORS: Shengshan Hu, Xiaogeng Liu, Yechao Zhang, Minghui Li, Leo Yu Zhang, Hai Jin, Libing Wu
HIGHLIGHT: In this paper, we propose adversarial makeup transfer GAN (AMT-GAN), a novel face protection method aiming at constructing adversarial face images that preserve stronger black-box transferability and better visual quality simultaneously.

593, TITLE: PONI: Potential Functions for ObjectGoal Navigation With Interaction-Free Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Ramakrishnan_PONI_Potential_Functions_for_ObjectGoal_Navigation_With_Interaction-Free_Learning_CVPR_2022_paper.html
AUTHORS: Santhosh Kumar Ramakrishnan, Devendra Singh Chaplot, Ziad Al-Halah, Jitendra Malik, Kristen Grauman
HIGHLIGHT: We propose Potential functions for ObjectGoal Navigation with Interaction-free learning (PONI), a modular approach that disentangles the skills of 'where to look?'

594, TITLE:		Noise Is Also Useful: Negative Correlation-Steered Latent Contrastive Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Yan_Noise_Is_Also_Useful_Negative_Correlation-Steered_Latent_Contrastive_Learning_CVPR_2022_paper.html
AUTHORS:		Jiexi Yan, Lei Luo, Chenghao Xu, Cheng Deng, Heng Huang
HIGHLIGHT:		However, they cannot guarantee the robust generalization of models due to the ignorance of useful information hidden in noisy data. To address this issue, we propose a new effective method named as LaCoL (Latent Contrastive Learning) to leverage the negative correlations from the noisy data.

595, TITLE:		Temporal Feature Alignment and Mutual Information Maximization for Video-Based Human Pose Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Temporal_Feature_Alignment_and_Mutual_Information_Maximization_for_Video-Based_Human_CVPR_2022_paper.html
AUTHORS:		Zhenguang Liu, Runyang Feng, Haoming Chen, Shuang Wu, Yixing Gao, Yunjun Gao, Xiang Wang
HIGHLIGHT:		More importantly, existing approaches build upon the straightforward pose estimation loss, which unfortunately cannot constrain the network to fully leverage useful information from neighboring frames. To tackle these problems, we present a novel hierarchical alignment framework, which leverages coarse-to-fine deformations to progressively update a neighboring frame to align with the current frame at the feature level.

596, TITLE:		Spatially-Adaptive Multilayer Selection for GAN Inversion and Editing
http://openaccess.thecvf.com//content/CVPR2022/html/Parmar_Spatially-Adaptive_Multilayer_Selection_for_GAN_Inversion_and_Editing_CVPR_2022_paper.html
AUTHORS:		Gaurav Parmar, Yijun Li, Jingwan Lu, Richard Zhang, Jun-Yan Zhu, Krishna Kumar Singh
HIGHLIGHT:		Existing GAN inversion and editing methods work well for aligned objects with a clean background, such as portraits and animal faces, but often struggle for more difficult categories with complex scene layouts and object occlusions, such as cars, animals, and outdoor images. We propose a new method to invert and edit such complex images in the latent space of GANs, such as StyleGAN2.

597, TITLE:		Self-Supervised Transformers for Unsupervised Object Discovery Using Normalized Cut
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Self-Supervised_Transformers_for_Unsupervised_Object_Discovery_Using_Normalized_Cut_CVPR_2022_paper.html
AUTHORS:		Yangtao Wang, Xi Shen, Shell Xu Hu, Yuan Yuan, James L. Crowley, Dominique Vaufreydaz
HIGHLIGHT:		In this paper, we show a graph-based method that uses the self-supervised transformer features to discover an object from an image.

598, TITLE:		Exploring Structure-Aware Transformer Over Interaction Proposals for Human-Object Interaction Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Exploring_Structure-Aware_Transformer_Over_Interaction_Proposals_for_Human-Object_Interaction_Detection_CVPR_2022_paper.html
AUTHORS:		Yong Zhang, Yingwei Pan, Ting Yao, Rui Huang, Tao Mei, Chang-Wen Chen
HIGHLIGHT:		In this work, we design a novel Transformer-style HOI detector, i.e., Structure-aware Transformer over Interaction Proposals (STIP), for HOI detection.

599, TITLE:		Towards Robust Adaptive Object Detection Under Noisy Annotations
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Towards_Robust_Adaptive_Object_Detection_Under_Noisy_Annotations_CVPR_2022_paper.html
AUTHORS:		Xinyu Liu, Wuyang Li, Qiushi Yang, Baopu Li, Yixuan Yuan
HIGHLIGHT:		Existing methods assume that the source domain labels are completely clean, yet large-scale datasets often contain error-prone annotations due to instance ambiguity, which may lead to a biased source distribution and severely degrade the performance of the domain adaptive detector de facto. In this paper, we represent the first effort to formulate noisy DAOD and propose a Noise Latent Transferability Exploration (NLTE) framework to address this issue.

600, TITLE:		Decoupled Multi-Task Learning With Cyclical Self-Regulation for Face Parsing
http://openaccess.thecvf.com//content/CVPR2022/html/Zheng_Decoupled_Multi-Task_Learning_With_Cyclical_Self-Regulation_for_Face_Parsing_CVPR_2022_paper.html
AUTHORS:		Qingping Zheng, Jiankang Deng, Zheng Zhu, Ying Li, Stefanos Zafeiriou
HIGHLIGHT:		This paper probes intrinsic factors behind typical failure cases (e.g spatial inconsistency and boundary confusion) produced by the existing state-of-the-art method in face parsing. To tackle these problems, we propose a novel Decoupled Multi-task Learning with Cyclical Self-Regulation (DML-CSR) for face parsing.

601, TITLE:		Pastiche Master: Exemplar-Based High-Resolution Portrait Style Transfer
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Pastiche_Master_Exemplar-Based_High-Resolution_Portrait_Style_Transfer_CVPR_2022_paper.html
AUTHORS:		Shuai Yang, Liming Jiang, Ziwei Liu, Chen Change Loy
HIGHLIGHT:		In this paper, we explore more challenging exemplar-based high-resolution portrait style transfer by introducing a novel DualStyleGAN with flexible control of dual styles of the original face domain and the extended artistic portrait domain.

602, TITLE:		Learning To Memorize Feature Hallucination for One-Shot Image Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Xie_Learning_To_Memorize_Feature_Hallucination_for_One-Shot_Image_Generation_CVPR_2022_paper.html
AUTHORS:		Yu Xie, Yanwei Fu, Ying Tai, Yun Cao, Junwei Zhu, Chengjie Wang

HIGHLIGHT: In this paper, we propose a novel model to explicitly learn and memorize reusable features that can help hallucinate novel category images.

603, TITLE: AUV-Net: Learning Aligned UV Maps for Texture Transfer and Synthesis
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_AUV-Net_Learning_Aligned_UV_Maps_for_Texture_Transfer_and_Synthesis_CVPR_2022_paper.html
AUTHORS: Zhiqin Chen, Kangxue Yin, Sanja Fidler
HIGHLIGHT: In this paper, we address the problem of texture representation for 3D shapes for the challenging and underexplored tasks of texture transfer and synthesis.

604, TITLE: Open-Vocabulary One-Stage Detection With Hierarchical Visual-Language Knowledge Distillation
http://openaccess.thecvf.com//content/CVPR2022/html/Ma_Open-Vocabulary_One-Stage_Detection_With_Hierarchical_Visual-Language_Knowledge_Distillation_CVPR_2022_paper.html
AUTHORS: Zongyang Ma, Guan Luo, Jin Gao, Liang Li, Yuxin Chen, Shaoru Wang, Congxuan Zhang, Weiming Hu
HIGHLIGHT: In this paper, we propose a hierarchical visual-language knowledge distillation method, i.e., HierKD, for open-vocabulary one-stage detection.

605, TITLE: Glass: Geometric Latent Augmentation for Shape Spaces
http://openaccess.thecvf.com//content/CVPR2022/html/Muralikrishnan_Glass_Geometric_Latent_Augmentation_for_Shape_Spaces_CVPR_2022_paper.html
AUTHORS: Sanjeev Muralikrishnan, Siddhartha Chaudhuri, Noam Aigerman, Vladimir G. Kim, Matthew Fisher, Niloy J. Mitra
HIGHLIGHT: We investigate the problem of training generative models on very sparse collections of 3D models.

606, TITLE: COAP: Compositional Articulated Occupancy of People
http://openaccess.thecvf.com//content/CVPR2022/html/Mihajlovic_COAP_Compositional_Articulated_Occupancy_of_People_CVPR_2022_paper.html
AUTHORS: Marko Mihajlovic, Shunsuke Saito, Aayush Bansal, Michael Zollhöfer, Siyu Tang
HIGHLIGHT: We present a novel neural implicit representation for articulated human bodies.

607, TITLE: Counterfactual Cycle-Consistent Learning for Instruction Following and Generation in Vision-Language Navigation
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Counterfactual_Cycle-Consistent_Learning_for_Instruction_Following_and_Generation_in_Vision-Language_CVPR_2022_paper.html
AUTHORS: Hanqing Wang, Wei Liang, Jianbing Shen, Luc Van Gool, Wenguan Wang
HIGHLIGHT: Here we describe an approach that learns the two tasks simultaneously and exploits their intrinsic correlations to boost the training of each: the follower judges whether the speaker-created instruction explains the original navigation route correctly, and vice versa.

608, TITLE: Evading the Simplicity Bias: Training a Diverse Set of Models Discovers Solutions With Superior OOD Generalization
http://openaccess.thecvf.com//content/CVPR2022/html/Teney_Evading_the_Simplicity_Bias_Training_a_Diverse_Set_of_Models_CVPR_2022_paper.html
AUTHORS: Damien Teney, Ehsan Abbasnejad, Simon Lucey, Anton van den Hengel
HIGHLIGHT: We demonstrate that the simplicity bias can be mitigated and OOD generalization improved.

609, TITLE: Assembly101: A Large-Scale Multi-View Video Dataset for Understanding Procedural Activities
http://openaccess.thecvf.com//content/CVPR2022/html/Sener_Assembly101_A_Large-Scale_Multi-View_Video_Dataset_for_Understanding_Procedural_Activities_CVPR_2022_paper.html
AUTHORS: Fadime Sener, Dibyadip Chatterjee, Daniel Shelepov, Kun He, Dipika Singhania, Robert Wang, Angela Yao
HIGHLIGHT: Assembly101 is a new procedural activity dataset featuring 4321 videos of people assembling and disassembling 101 "take-apart" toy vehicles.

610, TITLE: Deterministic Point Cloud Registration via Novel Transformation Decomposition
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Deterministic_Point_Cloud_Registration_via_Novel_Transformation_Decomposition_CVPR_2022_paper.html
AUTHORS: Wen Chen, Haoang Li, Qiang Nie, Yun-Hui Liu
HIGHLIGHT: Given a set of putative 3D-3D point correspondences, we aim to remove outliers and estimate rigid transformation with 6 degrees of freedom (DOF).

611, TITLE: Motion-Adjustable Neural Implicit Video Representation
http://openaccess.thecvf.com//content/CVPR2022/html/Mai_Motion-Adjustable_Neural_Implicit_Video_Representation_CVPR_2022_paper.html
AUTHORS: Long Mai, Feng Liu
HIGHLIGHT: By exploiting the relation between the phase information in sinusoidal functions and their displacements, we incorporate into the conventional image-based INR model a phase-varying positional encoding module, and couple it with a phase-shift generation module that determines the phase-shift values at each frame.

612, TITLE: Neural Prior for Trajectory Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Neural_Prior_for_Trajectory_Estimation_CVPR_2022_paper.html

AUTHORS: Chaoyang Wang, Xueqian Li, Jhony Kaesemodel Pontes, Simon Lucey
HIGHLIGHT: Here, we propose a neural trajectory prior to capture continuous spatio-temporal information without the need for offline data.

613, TITLE: DPICT: Deep Progressive Image Compression Using Trit-Planes
http://openaccess.thecvf.com//content/CVPR2022/html/Lee_DPICT_Deep_Progressive_Image_Compression_Using_Trit-Planes_CVPR_2022_paper.html
AUTHORS: Jae-Han Lee, Seungmin Jeon, Kwang Pyo Choi, Youngo Park, Chang-Su Kim
HIGHLIGHT: We propose the deep progressive image compression using trit-planes (DPICT) algorithm, which is the first learning-based codec supporting fine granular scalability (FGS).

614, TITLE: Rethinking Depth Estimation for Multi-View Stereo: A Unified Representation
http://openaccess.thecvf.com//content/CVPR2022/html/Peng_Rethinking_Depth_Estimation_for_Multi-View_Stereo_A_Unified_Representation_CVPR_2022_paper.html
AUTHORS: Rui Peng, Rongjie Wang, Zhenyu Wang, Yawen Lai, Ronggang Wang
HIGHLIGHT: In this paper, we propose a novel representation, termed Unification, to unify the advantages of regression and classification.

615, TITLE: Long-Tailed Recognition via Weight Balancing
http://openaccess.thecvf.com//content/CVPR2022/html/Alshammari_Long-Tailed_Recognition_via_Weight_Balancing_CVPR_2022_paper.html
AUTHORS: Shaden Alshammari, Yu-Xiong Wang, Deva Ramanan, Shu Kong
HIGHLIGHT: Surprisingly, weight decay, although underexplored in LTR, significantly improves over prior work. Therefore, we adopt a two-stage training paradigm and propose a simple approach to LTR: (1) learning features using the cross-entropy loss by tuning weight decay, and (2) learning classifiers using class-balanced loss by tuning weight decay and MaxNorm.

616, TITLE: Text to Image Generation With Semantic-Spatial Aware GAN
http://openaccess.thecvf.com//content/CVPR2022/html/Liao_Text_to_Image_Generation_With_Semantic-Spatial_Aware_GAN_CVPR_2022_paper.html
AUTHORS: Wentong Liao, Kai Hu, Michael Ying Yang, Bodo Rosenhahn
HIGHLIGHT: A close inspection of their generated images reveals a major limitation: even though the generated image holistically matches the description, individual image regions or parts of somethings are often not recognizable or consistent with words in the sentence, e.g. "a white crown". To address this problem, we propose a novel framework Semantic-Spatial Aware GAN for synthesizing images from input text.

617, TITLE: The Norm Must Go On: Dynamic Unsupervised Domain Adaptation by Normalization
http://openaccess.thecvf.com//content/CVPR2022/html/Mirza_The_Norm_Must_Go_On_Dynamic_Unsupervised_Domain_Adaptation_by_CVPR_2022_paper.html
AUTHORS: M. Jehanzeb Mirza, Jakub Micorek, Horst Possegger, Horst Bischof
HIGHLIGHT: This can be a hurdle in fields which require continuous dynamic adaptation or suffer from scarcity of data, e.g. autonomous driving in challenging weather conditions. To address this problem of continuous adaptation to distribution shifts, we propose Dynamic Unsupervised Adaptation (DUA).

618, TITLE: ShapeFormer: Transformer-Based Shape Completion via Sparse Representation
http://openaccess.thecvf.com//content/CVPR2022/html/Yan_ShapeFormer_Transformer-Based_Shape_Completion_via_Sparse_Representation_CVPR_2022_paper.html
AUTHORS: Xingguang Yan, Liqiang Lin, Niloy J. Mitra, Dani Lischinski, Daniel Cohen-Or, Hui Huang
HIGHLIGHT: We present ShapeFormer, a transformer-based network that produces a distribution of object completions, conditioned on incomplete, and possibly noisy, point clouds.

619, TITLE: PixMix: Dreamlike Pictures Comprehensively Improve Safety Measures
http://openaccess.thecvf.com//content/CVPR2022/html/Hendrycks_PixMix_Dreamlike_Pictures_Comprehensively_Improve_Safety_Measures_CVPR_2022_paper.html
AUTHORS: Dan Hendrycks, Andy Zou, Mantas Mazeika, Leonard Tang, Bo Li, Dawn Song, Jacob Steinhardt
HIGHLIGHT: Similarly, strong data augmentation and regularization techniques often improve OOD robustness but harm anomaly detection, raising the question of whether a Pareto improvement on all existing safety measures is possible. To meet this challenge, we design a new data augmentation strategy utilizing the natural structural complexity of pictures such as fractals, which outperforms numerous baselines, is near Pareto-optimal, and roundly improves safety measures.

620, TITLE: Eigencontours: Novel Contour Descriptors Based on Low-Rank Approximation
http://openaccess.thecvf.com//content/CVPR2022/html/Park_Eigencontours_Novel_Contour_Descriptors_Based_on_Low-Rank_Approximation_CVPR_2022_paper.html
AUTHORS: Wonhui Park, Dongkwon Jin, Chang-Su Kim
HIGHLIGHT: Novel contour descriptors, called eigencontours, based on low-rank approximation are proposed in this paper.

621, TITLE: Generalizable Cross-Modality Medical Image Segmentation via Style Augmentation and Dual Normalization
http://openaccess.thecvf.com//content/CVPR2022/html/Zhou_Generalizable_Cross-Modality_Medical_Image_Segmentation_via_Style_Augmentation_and_Dual_CVPR_2022_paper.html
AUTHORS: Ziqi Zhou, Lei Qi, Xin Yang, Dong Ni, Yinghuan Shi

HIGHLIGHT: This setting, namely generalizable cross-modality segmentation, owning its clinical potential, is much more challenging than other related settings, e.g., domain adaptation. To achieve this goal, we in this paper propose a novel dual-normalization model by leveraging the augmented source-similar and source-dissimilar images during our generalizable segmentation.

622, TITLE: Learning Optical Flow With Kernel Patch Attention
http://openaccess.thecvf.com//content/CVPR2022/html/Luo_Learning_Optical_Flow_With_Kernel_Patch_Attention_CVPR_2022_paper.html
AUTHORS: Ao Luo, Fan Yang, Xin Li, Shuaicheng Liu
HIGHLIGHT: In this paper, we introduce a novel approach, called kernel patch attention (KPA), to better resolve the ambiguity in dense matching by explicitly taking the local context relations into consideration.

623, TITLE: Learning To Prompt for Open-Vocabulary Object Detection With Vision-Language Model
http://openaccess.thecvf.com//content/CVPR2022/html/Du_Learning_To_Prompt_for_Open-Vocabulary_Object_Detection_With_Vision-Language_Model_CVPR_2022_paper.html
AUTHORS: Yu Du, Fangyun Wei, Zihe Zhang, Miaojing Shi, Yue Gao, Guoqi Li
HIGHLIGHT: In this paper, we introduce a novel method, detection prompt (DetPro), to learn continuous prompt representations for open-vocabulary object detection based on the pre-trained vision-language model.

624, TITLE: TimeReplayer: Unlocking the Potential of Event Cameras for Video Interpolation
http://openaccess.thecvf.com//content/CVPR2022/html/He_TimeReplayer_Unlocking_the_Potential_of_Event_Cameras_for_Video_Interpolation_CVPR_2022_paper.html
AUTHORS: Weihua He, Kaichao You, Zhendong Qiao, Xu Jia, Ziyang Zhang, Wenhui Wang, Huchuan Lu, Yaoyuan Wang, Jianxing Liao
HIGHLIGHT: To fully unlock the potential of event cameras, this paper proposes a novel TimeReplayer algorithm to interpolate videos captured by commodity cameras with events.

625, TITLE: General Incremental Learning With Domain-Aware Categorical Representations
http://openaccess.thecvf.com//content/CVPR2022/html/Xie_General_Incremental_Learning_With_Domain-Aware_Categorical_Representations_CVPR_2022_paper.html
AUTHORS: Jiangwei Xie, Shipeng Yan, Xuming He
HIGHLIGHT: In this work, we consider a general and yet under-explored incremental learning problem in which both the class distribution and class-specific domain distribution change over time.

626, TITLE: Interactive Segmentation and Visualization for Tiny Objects in Multi-Megapixel Images
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_Interactive_Segmentation_and_Visualization_for_Tiny_Objects_in_Multi-Megapixel_Images_CVPR_2022_paper.html
AUTHORS: Chengyuan Xu, Boning Dong, Noah Stier, Curtis McCully, D. Andrew Howell, Pradeep Sen, Tobias Höllerer
HIGHLIGHT: We introduce an interactive image segmentation and visualization framework for identifying, inspecting, and editing tiny objects (just a few pixels wide) in large multi-megapixel high-dynamic-range (HDR) images.

627, TITLE: ActiveZero: Mixed Domain Learning for Active Stereovision With Zero Annotation
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_ActiveZero_Mixed_Domain_Learning_for_Active_Stereovision_With_Zero_Annotation_CVPR_2022_paper.html
AUTHORS: Isabella Liu, Edward Yang, Jianyu Tao, Rui Chen, Xiaoshuai Zhang, Qing Ran, Zhu Liu, Hao Su
HIGHLIGHT: Since ground truth depth is readily available in the simulation domain but quite difficult to obtain in the real domain, we propose a method that leverages the best of both worlds.

628, TITLE: DearKD: Data-Efficient Early Knowledge Distillation for Vision Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_DearKD_Data-Efficient_Early_Knowledge_Distillation_for_Vision_Transformers_CVPR_2022_paper.html
AUTHORS: Xianing Chen, Qiong Cao, Yujie Zhong, Jing Zhang, Shenghua Gao, Dacheng Tao
HIGHLIGHT: In this work, we propose an early knowledge distillation framework, which is termed as DearKD, to improve the data-efficiency required by transformers.

629, TITLE: Global-Aware Registration of Less-Overlap RGB-D Scans
http://openaccess.thecvf.com//content/CVPR2022/html/Sun_Global-Aware_Registration_of_Less-Overlap_RGB-D_Scans_CVPR_2022_paper.html
AUTHORS: Che Sun, Yunde Jia, Yi Guo, Yuwei Wu
HIGHLIGHT: We propose a novel method of registering less-overlap RGB-D scans.

630, TITLE: RayMVSNet: Learning Ray-Based 1D Implicit Fields for Accurate Multi-View Stereo
http://openaccess.thecvf.com//content/CVPR2022/html/Xi_RayMVSNet_Learning_Ray-Based_1D_Implicit_Fields_for_Accurate_Multi-View_Stereo_CVPR_2022_paper.html
AUTHORS: Junhua Xi, Yifei Shi, Yijie Wang, Yulan Guo, Kai Xu
HIGHLIGHT: In particular, we propose RayMVSNet which learns sequential prediction of a 1D implicit field along each camera ray with the zero-crossing point indicating scene depth.

631, TITLE: ContrastMask: Contrastive Learning To Segment Every Thing

http://openaccess.thecvf.com//content/CVPR2022/html/Wang_ContrastMask_Contrastive_Learning_To_Segment_Every_Thing_CVPR_2022_paper.html
AUTHORS: Xuehui Wang, Kai Zhao, Ruixin Zhang, Shouhong Ding, Yan Wang, Wei Shen
HIGHLIGHT: Unlike previous methods that learn such models only on base categories, in this paper, we propose a new method, named ContrastMask, which learns a mask segmentation model on both base and novel categories under a unified pixel-level contrastive learning framework.

632, TITLE: Efficient Deep Embedded Subspace Clustering
http://openaccess.thecvf.com//content/CVPR2022/html/Cai_Efficient_Deep_Embedded_Subspace_Clustering_CVPR_2022_paper.html
AUTHORS: Jinyu Cai, Jicong Fan, Wenzhong Guo, Shiping Wang, Yunhe Zhang, Zhao Zhang
HIGHLIGHT: In this paper, we propose a new mechanism for deep clustering.

633, TITLE: Neural MoCon: Neural Motion Control for Physically Plausible Human Motion Capture
http://openaccess.thecvf.com//content/CVPR2022/html/Huang_Neural_MoCon_Neural_Motion_Control_for_Physically_Plausible_Human_Motion_CVPR_2022_paper.html
AUTHORS: Buzhen Huang, Liang Pan, Yuan Yang, Jingyi Ju, Yangang Wang
HIGHLIGHT: In this work, we focus on exploiting the high-precision and non-differentiable physics simulator to incorporate dynamical constraints in motion capture.

634, TITLE: Revisiting Temporal Alignment for Video Restoration
http://openaccess.thecvf.com//content/CVPR2022/html/Zhou_Revisiting_Temporal_Alignment_for_Video_Restoration_CVPR_2022_paper.html
AUTHORS: Kun Zhou, Wenbo Li, Liying Lu, Xiaoguang Han, Jiangbo Lu
HIGHLIGHT: In this work, we present a novel, generic iterative alignment module which employs a gradual refinement scheme for sub-alignments, yielding more accurate motion compensation.

635, TITLE: Scaling Vision Transformers to Gigapixel Images via Hierarchical Self-Supervised Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Scaling_Vision_Transformers_to_Gigapixel_Images_via_Hierarchical_Self-Supervised_Learning_CVPR_2022_paper.html
AUTHORS: Richard J. Chen, Chengkuan Chen, Yicong Li, Tiffany Y. Chen, Andrew D. Trister, Rahul G. Krishnan, Faisal Mahmood
HIGHLIGHT: We introduce a new ViT architecture called the Hierarchical Image Pyramid Transformer (HIPT), which leverages the natural hierarchical structure inherent in WSIs using two levels of self-supervised learning to learn high-resolution image representations.

636, TITLE: Neural Reflectance for Shape Recovery With Shadow Handling
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Neural_Reflectance_for_Shape_Recovery_With_Shadow_Handling_CVPR_2022_paper.html
AUTHORS: Junxuan Li, Hongdong Li
HIGHLIGHT: This paper aims at recovering the shape of a scene with unknown, non-Lambertian, and possibly spatially-varying surface materials.

637, TITLE: Rep-Net: Efficient On-Device Learning via Feature Reprogramming
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Rep-Net_Efficient_On-Device_Learning_via_Feature_Reprogramming_CVPR_2022_paper.html
AUTHORS: Li Yang, Adnan Siraj Rakin, Deliang Fan
HIGHLIGHT: To develop memory-efficient on-device transfer learning, in this work, we are the first to approach the concept of transfer learning from a new perspective of intermediate feature reprogramming of a pre-trained model (i.e., backbone).

638, TITLE: Surface Representation for Point Clouds
http://openaccess.thecvf.com//content/CVPR2022/html/Ran_Surface_Representation_for_Point_Clouds_CVPR_2022_paper.html
AUTHORS: Haoxi Ran, Jun Liu, Chengjie Wang
HIGHLIGHT: In this paper, we present RepSurf (representative surfaces), a novel representation of point clouds to explicitly depict the very local structure.

639, TITLE: Implicit Motion Handling for Video Camouflaged Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Cheng_Implicit_Motion_Handling_for_Video_Camouflaged_Object_Detection_CVPR_2022_paper.html
AUTHORS: Xuelian Cheng, Huan Xiong, Deng-Ping Fan, Yiran Zhong, Mehrtash Harandi, Tom Drummond, Zongyuan Ge
HIGHLIGHT: We propose a new video camouflaged object detection (VCOD) framework that can exploit both short-term dynamics and long-term temporal consistency to detect camouflaged objects from video frames.

640, TITLE: OVE6D: Object Viewpoint Encoding for Depth-Based 6D Object Pose Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Cai_OVE6D_Object_Viewpoint_Encoding_for_Depth-Based_6D_Object_Pose_Estimation_CVPR_2022_paper.html
AUTHORS: Dingding Cai, Janne Heikkilä, Esa Rahtu
HIGHLIGHT: This paper proposes a universal framework, called OVE6D, for model-based 6D object pose estimation from a single depth image and a target object mask.

641, TITLE:　　　　DeepLIIF: An Online Platform for Quantification of Clinical Pathology Slides
http://openaccess.thecvf.com//content/CVPR2022/html/Ghahremani_DeepLIIF_An_Online_Platform_for_Quantification_of_Clinical
_Pathology_Slides_CVPR_2022_paper.html
AUTHORS:　　　　Parmida Ghahremani, Joseph Marino, Ricardo Dodds, Saad Nadeem
HIGHLIGHT:　　　　We present DeepLIIF (https://deepliif.org), a first free online platform for efficient and reproducible IHC
scoring.

642, TITLE:　　　　Joint Video Summarization and Moment Localization by Cross-Task Sample Transfer
http://openaccess.thecvf.com//content/CVPR2022/html/Jiang_Joint_Video_Summarization_and_Moment_Localization_by_Cross-
Task_Sample_Transfer_CVPR_2022_paper.html
AUTHORS:　　　　Hao Jiang, Yadong Mu
HIGHLIGHT:　　　　Approximately, the video summary can be treated as a sparse, redundancy-free version of the video moments.
Inspired by this observation, we propose an importance Propagation based collaborative Teaching Network (iPTNet).

643, TITLE:　　　　WALT: Watch and Learn 2D Amodal Representation From Time-Lapse Imagery
http://openaccess.thecvf.com//content/CVPR2022/html/Reddy_WALT_Watch_and_Learn_2D_Amodal_Representation_From_Time-
Lapse_Imagery_CVPR_2022_paper.html
AUTHORS:　　　　N. Dinesh Reddy, Robert Tamburo, Srinivasa G. Narasimhan
HIGHLIGHT:　　　　In this work, we present the best of both the real and synthetic worlds for automatic occlusion supervision using
a large readily available source of data: time-lapse imagery from stationary webcams observing street intersections over weeks,
months, or even years.

644, TITLE:　　　　Learning With Twin Noisy Labels for Visible-Infrared Person Re-Identification
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Learning_With_Twin_Noisy_Labels_for_Visible-Infrared_Person_Re-
Identification_CVPR_2022_paper.html
AUTHORS:　　　　Mouxing Yang, Zhenyu Huang, Peng Hu, Taihao Li, Jiancheng Lv, Xi Peng
HIGHLIGHT:　　　　In this paper, we study an untouched problem in visible-infrared person re-identification (VI-ReID), namely,
Twin Noise Labels (TNL) which refers to as noisy annotation and correspondence.

645, TITLE:　　　　Optical Flow Estimation for Spiking Camera
http://openaccess.thecvf.com//content/CVPR2022/html/Hu_Optical_Flow_Estimation_for_Spiking_Camera_CVPR_2022_paper.html
AUTHORS:　　　　Liwen Hu, Rui Zhao, Ziluo Ding, Lei Ma, Boxin Shi, Ruiqin Xiong, Tiejun Huang
HIGHLIGHT:　　　　However, frame-based and event-based methods are not well suited to spike streams from the spiking camera
due to the different data modalities. To this end, we present, SCFlow, a tailored deep learning pipeline to estimate optical flow in high-
speed scenes from spike streams.

646, TITLE:　　　　MetaFormer Is Actually What You Need for Vision
http://openaccess.thecvf.com//content/CVPR2022/html/Yu_MetaFormer_Is_Actually_What_You_Need_for_Vision_CVPR_2022_pa
per.html
AUTHORS:　　　　Weihao Yu, Mi Luo, Pan Zhou, Chenyang Si, Yichen Zhou, Xinchao Wang, Jiashi Feng, Shuicheng Yan
HIGHLIGHT:　　　　To verify this, we deliberately replace the attention module in transformers with an embarrassingly simple
spatial pooling operator to conduct only basic token mixing. Surprisingly, we observe that the derived model, termed as PoolFormer,
achieves competitive performance on multiple computer vision tasks.

647, TITLE:　　　　GradViT: Gradient Inversion of Vision Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Hatamizadeh_GradViT_Gradient_Inversion_of_Vision_Transformers_CVPR
_2022_paper.html
AUTHORS:　　　　Ali Hatamizadeh, Hongxu Yin, Holger R. Roth, Wenqi Li, Jan Kautz, Daguang Xu, Pavlo Molchanov
HIGHLIGHT:　　　　In this work we demonstrate the vulnerability of vision transformers (ViTs) to gradient-based inversion attacks.

648, TITLE:　　　　Spatial-Temporal Space Hand-in-Hand: Spatial-Temporal Video Super-Resolution via Cycle-Projected Mutual
Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Hu_Spatial-Temporal_Space_Hand-in-Hand_Spatial-Temporal_Video_Super-
Resolution_via_Cycle-Projected_Mutual_Learning_CVPR_2022_paper.html
AUTHORS:　　　　Mengshun Hu, Kui Jiang, Liang Liao, Jing Xiao, Junjun Jiang, Zheng Wang
HIGHLIGHT:　　　　To this end, we propose a one-stage based Cycle-projected Mutual learning network (CycMu-Net) for ST-VSR,
which makes full use of spatial-temporal correlations via the mutual learning between S-VSR and T-VSR.

649, TITLE:　　　　InstaFormer: Instance-Aware Image-to-Image Translation With Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Kim_InstaFormer_Instance-Aware_Image-to-
Image_Translation_With_Transformer_CVPR_2022_paper.html
AUTHORS:　　　　Soohyun Kim, Jongbeom Baek, Jihye Park, Gyeongnyeon Kim, Seungryong Kim
HIGHLIGHT:　　　　We present a novel Transformer-based network architecture for instance-aware image-to-image translation,
dubbed InstaFormer, to effectively integrate global- and instance-level information.

650, TITLE:　　　　Revisiting Near/Remote Sensing With Geospatial Attention
http://openaccess.thecvf.com//content/CVPR2022/html/Workman_Revisiting_NearRemote_Sensing_With_Geospatial_Attention_CV
PR_2022_paper.html
AUTHORS:　　　　Scott Workman, M. Usman Rafique, Hunter Blanton, Nathan Jacobs

HIGHLIGHT: We propose an approach for computing geospatial attention that incorporates geometric features and the appearance of the overhead and ground-level imagery.

651, TITLE: Joint Global and Local Hierarchical Priors for Learned Image Compression
http://openaccess.thecvf.com//content/CVPR2022/html/Kim_Joint_Global_and_Local_Hierarchical_Priors_for_Learned_Image_Compression_CVPR_2022_paper.html
AUTHORS: Jun-Hyuk Kim, Byeongho Heo, Jong-Seok Lee
HIGHLIGHT: However, CNNs have a limitation in modeling long-range dependencies due to their nature of local connectivity, which can be a significant bottleneck in image compression where reducing spatial redundancy is a key point. To overcome this issue, we propose a novel entropy model called Information Transformer (Informer) that exploits both global and local information in a content-dependent manner using an attention mechanism.

652, TITLE: Knowledge Distillation via the Target-Aware Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Lin_Knowledge_Distillation_via_the_Target-Aware_Transformer_CVPR_2022_paper.html
AUTHORS: Sihao Lin, Hongwei Xie, Bing Wang, Kaicheng Yu, Xiaojun Chang, Xiaodan Liang, Gang Wang
HIGHLIGHT: This greatly undermines the underlying assumption of the one-to-one distillation approach. To this end, we propose a novel one-to-all spatial matching knowledge distillation approach.

653, TITLE: Recurring the Transformer for Video Action Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Recurring_the_Transformer_for_Video_Action_Recognition_CVPR_2022_paper.html
AUTHORS: Jiewen Yang, Xingbo Dong, Liujun Liu, Chao Zhang, Jiajun Shen, Dahai Yu
HIGHLIGHT: We introduce a novel Recurrent Vision Transformer (RViT) framework for spatial-temporal representation learning to achieve the video action recognition task.

654, TITLE: Subspace Adversarial Training
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Subspace_Adversarial_Training_CVPR_2022_paper.html
AUTHORS: Tao Li, Yingwen Wu, Sizhe Chen, Kun Fang, Xiaolin Huang
HIGHLIGHT: To control the growth of the gradient, we propose a new AT method, Subspace Adversarial Training (Sub-AT), which constrains AT in a carefully extracted subspace.

655, TITLE: 3D-VField: Adversarial Augmentation of Point Clouds for Domain Generalization in 3D Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Lehner_3D-VField_Adversarial_Augmentation_of_Point_Clouds_for_Domain_Generalization_in_CVPR_2022_paper.html
AUTHORS: Alexander Lehner, Stefano Gasperini, Alvaro Marcos-Ramiro, Michael Schmidt, Mohammad-Ali Nikouei Mahani, Nassir Navab, Benjamin Busam, Federico Tombari
HIGHLIGHT: In this work, we substantially improve the generalization of 3D object detectors to out-of-domain data by deforming point clouds during training.

656, TITLE: Image Segmentation Using Text and Image Prompts
http://openaccess.thecvf.com//content/CVPR2022/html/Luddecke_Image_Segmentation_Using_Text_and_Image_Prompts_CVPR_2022_paper.html
AUTHORS: Timo Lüddecke, Alexander Ecker
HIGHLIGHT: Here we propose a system that can generate image segmentations based on arbitrary prompts at test time.

657, TITLE: AutoMine: An Unmanned Mine Dataset
http://openaccess.thecvf.com//content/CVPR2022/html/Li_AutoMine_An_Unmanned_Mine_Dataset_CVPR_2022_paper.html
AUTHORS: Yuchen Li, Zixuan Li, Siyu Teng, Yu Zhang, Yuhang Zhou, Yuchang Zhu, Dongpu Cao, Bin Tian, Yunfeng Ai, Zhe Xuanyuan, Long Chen
HIGHLIGHT: Moreover, the open-pit mine is one of the typical representatives for them. Therefore, we introduce the Autonomous driving dataset on the Mining scene (AutoMine) for positioning and perception tasks in this paper.

658, TITLE: Neural Data-Dependent Transform for Learned Image Compression
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Neural_Data-Dependent_Transform_for_Learned_Image_Compression_CVPR_2022_paper.html
AUTHORS: Dezhao Wang, Wenhan Yang, Yueyu Hu, Jiaying Liu
HIGHLIGHT: Learned image compression has achieved great success due to its excellent modeling capacity, but seldom further considers the Rate-Distortion Optimization (RDO) of each input image. To explore this potential in the learned codec, we make the first attempt to build a neural data-dependent transform and introduce a continuous online mode decision mechanism to jointly optimize the coding efficiency for each individual image.

659, TITLE: Background Activation Suppression for Weakly Supervised Object Localization
http://openaccess.thecvf.com//content/CVPR2022/html/Wu_Background_Activation_Suppression_for_Weakly_Supervised_Object_Localization_CVPR_2022_paper.html
AUTHORS: Pingyu Wu, Wei Zhai, Yang Cao
HIGHLIGHT: In this paper, we propose a Background Activation Suppression (BAS) method.

660, TITLE: How Many Observations Are Enough? Knowledge Distillation for Trajectory Forecasting

http://openaccess.thecvf.com//content/CVPR2022/html/Monti_How_Many_Observations_Are_Enough_Knowledge_Distillation_for_Trajectory_Forecasting_CVPR_2022_paper.html
AUTHORS: Alessio Monti, Angelo Porrello, Simone Calderara, Pasquale Coscia, Lamberto Ballan, Rita Cucchiara
HIGHLIGHT: On this account, the model would be fed with corrupted and noisy input data, thus fatally affecting its prediction performance. In this regard, we focus on delivering accurate predictions when only a few input observations are used, thus potentially lowering the risks associated with automatic perception.

661, TITLE: Evaluation-Oriented Knowledge Distillation for Deep Face Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Huang_Evaluation-Oriented_Knowledge_Distillation_for_Deep_Face_Recognition_CVPR_2022_paper.html
AUTHORS: Yuge Huang, Jiaxiang Wu, Xingkun Xu, Shouhong Ding
HIGHLIGHT: Inspired by the ultimate goal of KD methods, we propose a novel Evaluation oriented KD method (EKD) for deep face recognition to directly reduce the performance gap between the teacher and student models during training.

662, TITLE: Improving Subgraph Recognition With Variational Graph Information Bottleneck
http://openaccess.thecvf.com//content/CVPR2022/html/Yu_Improving_Subgraph_Recognition_With_Variational_Graph_Information_Bottleneck_CVPR_2022_paper.html
AUTHORS: Junchi Yu, Jie Cao, Ran He
HIGHLIGHT: However, GIB suffers from training instability and degenerated results due to its intrinsic optimization process. To tackle these issues, we reformulate the subgraph recognition problem into two steps: graph perturbation and subgraph selection, leading to a novel Variational Graph Information Bottleneck (VGIB) framework

663, TITLE: Slot-VPS: Object-Centric Representation Learning for Video Panoptic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Zhou_Slot-VPS_Object-Centric_Representation_Learning_for_Video_Panoptic_Segmentation_CVPR_2022_paper.html
AUTHORS: Yi Zhou, Hui Zhang, Hana Lee, Shuyang Sun, Pingjun Li, Yangguang Zhu, ByungIn Yoo, Xiaojuan Qi, Jae-Joon Han
HIGHLIGHT: However, this divide-and-conquer strategy requires complex post-processing in both spatial and temporal domains and is vulnerable to failures from surrogate tasks. In this paper, inspired by object-centric learning which learns compact and robust object representations, we present Slot-VPS, the first end-to-end framework for this task.

664, TITLE: Motion-From-Blur: 3D Shape and Motion Estimation of Motion-Blurred Objects in Videos
http://openaccess.thecvf.com//content/CVPR2022/html/Rozumnyi_Motion-From-Blur_3D_Shape_and_Motion_Estimation_of_Motion-Blurred_Objects_in_CVPR_2022_paper.html
AUTHORS: Denys Rozumnyi, Martin R. Oswald, Vittorio Ferrari, Marc Pollefeys
HIGHLIGHT: We propose a method for jointly estimating the 3D motion, 3D shape, and appearance of highly motion-blurred objects from a video.

665, TITLE: Efficient Video Instance Segmentation via Tracklet Query and Proposal
http://openaccess.thecvf.com//content/CVPR2022/html/Wu_Efficient_Video_Instance_Segmentation_via_Tracklet_Query_and_Proposal_CVPR_2022_paper.html
AUTHORS: Jialian Wu, Sudhir Yarram, Hui Liang, Tian Lan, Junsong Yuan, Jayan Eledath, Gérard Medioni
HIGHLIGHT: This paper proposes EfficientVIS, a fully end-to-end framework with efficient training and inference.

666, TITLE: Synthetic Generation of Face Videos With Plethysmograph Physiology
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Synthetic_Generation_of_Face_Videos_With_Plethysmograph_Physiology_CVPR_2022_paper.html
AUTHORS: Zhen Wang, Yunhao Ba, Pradyumna Chari, Oyku Deniz Bozkurt, Gianna Brown, Parth Patwa, Niranjan Vaddi, Laleh Jalilian, Achuta Kadambi
HIGHLIGHT: This paper proposes a scalable biophysical learning based method to generate physio-realistic synthetic rPPG videos given any reference image and target rPPG signal and shows that it could further improve the state-of-the-art physiological measurement and reduce the bias among different groups.

667, TITLE: TransRAC: Encoding Multi-Scale Temporal Correlation With Transformers for Repetitive Action Counting
http://openaccess.thecvf.com//content/CVPR2022/html/Hu_TransRAC_Encoding_Multi-Scale_Temporal_Correlation_With_Transformers_for_Repetitive_Action_CVPR_2022_paper.html
AUTHORS: Huazhang Hu, Sixun Dong, Yiqun Zhao, Dongze Lian, Zhengxin Li, Shenghua Gao
HIGHLIGHT: In the data-driven era, the degradation of such generalization capability is mainly attributed to the lack of long video datasets. To complement this margin, we introduce a new large-scale repetitive action counting dataset covering a wide variety of video lengths, along with more realistic situations where action interruption or action inconsistencies occur in the video.

668, TITLE: Hallucinated Neural Radiance Fields in the Wild
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Hallucinated_Neural_Radiance_Fields_in_the_Wild_CVPR_2022_paper.html
AUTHORS: Xingyu Chen, Qi Zhang, Xiaoyu Li, Yue Chen, Ying Feng, Xuan Wang, Jue Wang
HIGHLIGHT: Existing solutions adopt NeRF with a controllable appearance embedding to render novel views under various conditions, but they cannot render view-consistent images with an unseen appearance. To solve this problem, we present an end-to-end framework for constructing a hallucinated NeRF, dubbed as Ha-NeRF.

669, TITLE: NeuralHDHair: Automatic High-Fidelity Hair Modeling From a Single Image Using Implicit Neural Representations
http://openaccess.thecvf.com//content/CVPR2022/html/Wu_NeuralHDHair_Automatic_High-Fidelity_Hair_Modeling_From_a_Single_Image_Using_CVPR_2022_paper.html
AUTHORS: Keyu Wu, Yifan Ye, Lingchen Yang, Hongbo Fu, Kun Zhou, Youyi Zheng
HIGHLIGHT: In this paper, we introduce NeuralHDHair, a flexible, fully automatic system for modeling high-fidelity hair from a single image.

670, TITLE: The Two Dimensions of Worst-Case Training and Their Integrated Effect for Out-of-Domain Generalization
http://openaccess.thecvf.com//content/CVPR2022/html/Huang_The_Two_Dimensions_of_Worst-Case_Training_and_Their_Integrated_Effect_CVPR_2022_paper.html
AUTHORS: Zeyi Huang, Haohan Wang, Dong Huang, Yong Jae Lee, Eric P. Xing
HIGHLIGHT: Existing literature discussing this "hard-to-learn" concept are mainly expanded either along the dimension of the samples or the dimension of the features. In this paper, we aim to introduce a simple view merging these two dimensions, leading to a new, simple yet effective, heuristic to train machine learning models by emphasizing the worst-cases on both the sample and the feature dimensions.

671, TITLE: Global Tracking Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Zhou_Global_Tracking_Transformers_CVPR_2022_paper.html
AUTHORS: Xingyi Zhou, Tianwei Yin, Vladlen Koltun, Philipp Krähenbühl
HIGHLIGHT: We present a novel transformer-based architecture for global multi-object tracking.

672, TITLE: Backdoor Attacks on Self-Supervised Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Saha_Backdoor_Attacks_on_Self-Supervised_Learning_CVPR_2022_paper.html
AUTHORS: Aniruddha Saha, Ajinkya Tejankar, Soroush Abbasi Koohpayegani, Hamed Pirsiavash
HIGHLIGHT: Backdoor attacks have been studied extensively in supervised learning and to the best of our knowledge, we are the first to study them for self-supervised learning.

673, TITLE: Multimodal Token Fusion for Vision Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Multimodal_Token_Fusion_for_Vision_Transformers_CVPR_2022_paper.html
AUTHORS: Yikai Wang, Xinghao Chen, Lele Cao, Wenbing Huang, Fuchun Sun, Yunhe Wang
HIGHLIGHT: In this paper, we propose a multimodal token fusion method (TokenFusion), tailored for transformer-based vision tasks.

674, TITLE: Exploring Frequency Adversarial Attacks for Face Forgery Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Jia_Exploring_Frequency_Adversarial_Attacks_for_Face_Forgery_Detection_CVPR_2022_paper.html
AUTHORS: Shuai Jia, Chao Ma, Taiping Yao, Bangjie Yin, Shouhong Ding, Xiaokang Yang
HIGHLIGHT: In this paper, instead of injecting adversarial perturbations into the spatial domain, we propose a frequency adversarial attack method against face forgery detectors.

675, TITLE: GMFlow: Learning Optical Flow via Global Matching
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_GMFlow_Learning_Optical_Flow_via_Global_Matching_CVPR_2022_paper.html
AUTHORS: Haofei Xu, Jing Zhang, Jianfei Cai, Hamid Rezatofighi, Dacheng Tao
HIGHLIGHT: Specifically, we propose a GMFlow framework, which consists of three main components: a customized Transformer for feature enhancement, a correlation and softmax layer for global feature matching, and a self-attention layer for flow propagation.

676, TITLE: Learning Hierarchical Cross-Modal Association for Co-Speech Gesture Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Learning_Hierarchical_Cross-Modal_Association_for_Co-Speech_Gesture_Generation_CVPR_2022_paper.html
AUTHORS: Xian Liu, Qianyi Wu, Hang Zhou, Yinghao Xu, Rui Qian, Xinyi Lin, Xiaowei Zhou, Wayne Wu, Bo Dai, Bolei Zhou
HIGHLIGHT: To fully utilize the rich connections between speech audio and human gestures, we propose a novel framework named Hierarchical Audio-to-Gesture (HA2G) for co-speech gesture generation.

677, TITLE: FLAVA: A Foundational Language and Vision Alignment Model
http://openaccess.thecvf.com//content/CVPR2022/html/Singh_FLAVA_A_Foundational_Language_and_Vision_Alignment_Model_CVPR_2022_paper.html
AUTHORS: Amanpreet Singh, Ronghang Hu, Vedanuj Goswami, Guillaume Couairon, Wojciech Galuba, Marcus Rohrbach, Douwe Kiela
HIGHLIGHT: A promising direction would be to use a single holistic universal model, as a "foundation", that targets all modalities at once---a true vision and language foundation model should be good at vision tasks, language tasks, and cross- and multi-modal vision and language tasks. We introduce FLAVA as such a model and demonstrate impressive performance on a wide range of 35 tasks spanning these target modalities.

678, TITLE: Signing at Scale: Learning to Co-Articulate Signs for Large-Scale Photo-Realistic Sign Language Production

http://openaccess.thecvf.com//content/CVPR2022/html/Saunders_Signing_at_Scale_Learning_to_Co-Articulate_Signs_for_Large-Scale_Photo-Realistic_CVPR_2022_paper.html
AUTHORS: Ben Saunders, Necati Cihan Camgoz, Richard Bowden
HIGHLIGHT: In this work, we tackle large-scale SLP by learning to co-articulate between dictionary signs, a method capable of producing smooth signing while scaling to unconstrained domains of discourse.

679, TITLE: Explore Spatio-Temporal Aggregation for Insubstantial Object Detection: Benchmark Dataset and Baseline
http://openaccess.thecvf.com//content/CVPR2022/html/Zhou_Explore_Spatio-Temporal_Aggregation_for_Insubstantial_Object_Detection_Benchmark_Dataset_and_CVPR_2022_paper.html
AUTHORS: Kailai Zhou, Yibo Wang, Tao Lv, Yunqian Li, Linsen Chen, Qiu Shen, Xun Cao
HIGHLIGHT: We endeavor on a rarely explored task named Insubstan-tial Object Detection (IOD), which aims to localize the object with following characteristics: (1) amorphous shape with indistinct boundary; (2) similarity to surroundings; (3) absence in color. Accordingly, it is far more challenging to distinguish insubstantial objects in a single static frame and the collaborative representation of spatial and tempo-ral information is crucial.

680, TITLE: OCSampler: Compressing Videos to One Clip With Single-Step Sampling
http://openaccess.thecvf.com//content/CVPR2022/html/Lin_OCSampler_Compressing_Videos_to_One_Clip_With_Single-Step_Sampling_CVPR_2022_paper.html
AUTHORS: Jintao Lin, Haodong Duan, Kai Chen, Dahua Lin, Limin Wang
HIGHLIGHT: There have been works that formulate frame sampling as a sequential decision task by selecting frames one by one according to their importance. In this paper, we present a more efficient framework named OCSampler, which explores such a representation with one short clip.

681, TITLE: Learning Bayesian Sparse Networks With Full Experience Replay for Continual Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Yan_Learning_Bayesian_Sparse_Networks_With_Full_Experience_Replay_for_Continual_CVPR_2022_paper.html
AUTHORS: Qingsen Yan, Dong Gong, Yuhang Liu, Anton van den Hengel, Javen Qinfeng Shi
HIGHLIGHT: Despite their performance, they still suffer from interference across tasks which leads to catastrophic forgetting. To ameliorate this problem, we propose to only activate and select sparse neurons for learning current and past tasks at any stage.

682, TITLE: Graph-Based Spatial Transformer With Memory Replay for Multi-Future Pedestrian Trajectory Prediction
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Graph-Based_Spatial_Transformer_With_Memory_Replay_for_Multi-Future_Pedestrian_Trajectory_CVPR_2022_paper.html
AUTHORS: Lihuan Li, Maurice Pagnucco, Yang Song
HIGHLIGHT: Our model aims to forecast multiple paths based on a historical trajectory by modeling multi-scale graph-based spatial transformers combined with a trajectory smoothing algorithm named "Memory Replay" utilizing a memory graph.

683, TITLE: Scanline Homographies for Rolling-Shutter Plane Absolute Pose
http://openaccess.thecvf.com//content/CVPR2022/html/Bai_Scanline_Homographies_for_Rolling-Shutter_Plane_Absolute_Pose_CVPR_2022_paper.html
AUTHORS: Fang Bai, Agniva Sengupta, Adrien Bartoli
HIGHLIGHT: In this work, we give a solution to the absolute pose problem free of motion assumptions.

684, TITLE: TableFormer: Table Structure Understanding With Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Nassar_TableFormer_Table_Structure_Understanding_With_Transformers_CVPR_2022_paper.html
AUTHORS: Ahmed Nassar, Nikolaos Livathinos, Maksym Lysak, Peter Staar
HIGHLIGHT: In this paper, we present a new table-structure identification model.

685, TITLE: Exemplar-Based Pattern Synthesis With Implicit Periodic Field Network
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Exemplar-Based_Pattern_Synthesis_With_Implicit_Periodic_Field_Network_CVPR_2022_paper.html
AUTHORS: Haiwei Chen, Jiayi Liu, Weikai Chen, Shichen Liu, Yajie Zhao
HIGHLIGHT: In this paper, we propose an exemplar-based visual pattern synthesis framework that aims to model the inner statistics of visual patterns and generate new, versatile patterns that meet the aforementioned requirements.

686, TITLE: Grounded Language-Image Pre-Training
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Grounded_Language-Image_Pre-Training_CVPR_2022_paper.html
AUTHORS: Liunian Harold Li, Pengchuan Zhang, Haotian Zhang, Jianwei Yang, Chunyuan Li, Yiwu Zhong, Lijuan Wang, Lu Yuan, Lei Zhang, Jenq-Neng Hwang, Kai-Wei Chang, Jianfeng Gao
HIGHLIGHT: This paper presents a grounded language-image pre-training (GLIP) model for learning object-level, language-aware, and semantic-rich visual representations.

687, TITLE: Spectral Unsupervised Domain Adaptation for Visual Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Spectral_Unsupervised_Domain_Adaptation_for_Visual_Recognition_CVPR_2022_paper.html
AUTHORS: Jingyi Zhang, Jiaxing Huang, Zichen Tian, Shijian Lu
HIGHLIGHT: We propose Spectral UDA (SUDA), an effective and efficient UDA technique that works in the spectral space and can generalize across different visual recognition tasks.

688, TITLE: AdaInt: Learning Adaptive Intervals for 3D Lookup Tables on Real-Time Image Enhancement
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_AdaInt_Learning_Adaptive_Intervals_for_3D_Lookup_Tables_on_Real-Time_CVPR_2022_paper.html
AUTHORS: Canqian Yang, Meiguang Jin, Xu Jia, Yi Xu, Ying Chen
HIGHLIGHT: They adopt a sub-optimal uniform sampling point allocation, limiting the expressiveness of the learned LUTs since the (tri-)linear interpolation between uniform sampling points in the LUT transform might fail to model local non-linearities of the color transform. Focusing on this problem, we present AdaInt (Adaptive Intervals Learning), a novel mechanism to achieve a more flexible sampling point allocation by adaptively learning the non-uniform sampling intervals in the 3D color space.

689, TITLE: PatchFormer: An Efficient Point Transformer With Patch Attention
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_PatchFormer_An_Efficient_Point_Transformer_With_Patch_Attention_CVPR_2022_paper.html
AUTHORS: Cheng Zhang, Haocheng Wan, Xinyi Shen, Zizhao Wu
HIGHLIGHT: However, existing point Transformers are computationally expensive since they need to generate a large attention map, which has quadratic complexity (both in space and time) with respect to input size. To solve this shortcoming, we introduce patch-attention (PAT) to adaptively learn a much smaller set of bases upon which the attention maps are computed.

690, TITLE: Recurrent Glimpse-Based Decoder for Detection With Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Recurrent_Glimpse-Based_Decoder_for_Detection_With_Transformer_CVPR_2022_paper.html
AUTHORS: Zhe Chen, Jing Zhang, Dacheng Tao
HIGHLIGHT: Alternative to existing studies that mainly develop advanced feature or embedding designs to tackle the training issue, we point out that the Region-of-Interest (RoI) based detection refinement can easily help mitigate the difficulty of training for DETR methods. Based on this, we introduce a novel REcurrent Glimpse-based decOder (REGO) in this paper.

691, TITLE: Generating 3D Bio-Printable Patches Using Wound Segmentation and Reconstruction To Treat Diabetic Foot Ulcers
http://openaccess.thecvf.com//content/CVPR2022/html/Chae_Generating_3D_Bio-Printable_Patches_Using_Wound_Segmentation_and_Reconstruction_To_CVPR_2022_paper.html
AUTHORS: Han Joo Chae, Seunghwan Lee, Hyewon Son, Seungyeob Han, Taebin Lim
HIGHLIGHT: We introduce AiD Regen, a novel system that generates 3D wound models combining 2D semantic segmentation with 3D reconstruction so that they can be printed via 3D bio-printers during the surgery to treat diabetic foot ulcers (DFUs).

692, TITLE: SimMIM: A Simple Framework for Masked Image Modeling
http://openaccess.thecvf.com//content/CVPR2022/html/Xie_SimMIM_A_Simple_Framework_for_Masked_Image_Modeling_CVPR_2022_paper.html
AUTHORS: Zhenda Xie, Zheng Zhang, Yue Cao, Yutong Lin, Jianmin Bao, Zhuliang Yao, Qi Dai, Han Hu
HIGHLIGHT: This paper presents SimMIM, a simple framework for masked image modeling.

693, TITLE: OmniFusion: 360 Monocular Depth Estimation via Geometry-Aware Fusion
http://openaccess.thecvf.com//content/CVPR2022/html/Li_OmniFusion_360_Monocular_Depth_Estimation_via_Geometry-Aware_Fusion_CVPR_2022_paper.html
AUTHORS: Yuyan Li, Yuliang Guo, Zhixin Yan, Xinyu Huang, Ye Duan, Liu Ren
HIGHLIGHT: In this paper, we propose a 360 monocular depth estimation pipeline, OmniFusion, to tackle the spherical distortion issue.

694, TITLE: Label Matching Semi-Supervised Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Label_Matching_Semi-Supervised_Object_Detection_CVPR_2022_paper.html
AUTHORS: Binbin Chen, Weijie Chen, Shicai Yang, Yunyi Xuan, Jie Song, Di Xie, Shiliang Pu, Mingli Song, Yueting Zhuang
HIGHLIGHT: Despite the promising results, the label mismatch problem is not yet fully explored in the previous works, leading to severe confirmation bias during self-training. In this paper, we delve into this problem and propose a simple yet effective LabelMatch framework from two different yet complementary perspectives, i.e., distribution-level and instance-level.

695, TITLE: RegionCLIP: Region-Based Language-Image Pretraining
http://openaccess.thecvf.com//content/CVPR2022/html/Zhong_RegionCLIP_Region-Based_Language-Image_Pretraining_CVPR_2022_paper.html
AUTHORS: Yiwu Zhong, Jianwei Yang, Pengchuan Zhang, Chunyuan Li, Noel Codella, Liunian Harold Li, Luowei Zhou, Xiyang Dai, Lu Yuan, Yin Li, Jianfeng Gao
HIGHLIGHT: To mitigate this issue, we propose a new method called RegionCLIP that significantly extends CLIP to learn region-level visual representations, thus enabling fine-grained alignment between image regions and textual concepts.

696, TITLE: Video Frame Interpolation Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Shi_Video_Frame_Interpolation_Transformer_CVPR_2022_paper.html
AUTHORS: Zhihao Shi, Xiangyu Xu, Xiaohong Liu, Jun Chen, Ming-Hsuan Yang
HIGHLIGHT: Existing methods for video interpolation heavily rely on deep convolution neural networks, and thus suffer from their intrinsic limitations, such as content-agnostic kernel weights and restricted receptive field. To address these issues, we propose a

Transformer-based video interpolation framework that allows content-aware aggregation weights and considers long-range dependencies with the self-attention operations.

697, TITLE: An MIL-Derived Transformer for Weakly Supervised Point Cloud Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_An_MIL-Derived_Transformer_for_Weakly_Supervised_Point_Cloud_Segmentation_CVPR_2022_paper.html
AUTHORS: Cheng-Kun Yang, Ji-Jia Wu, Kai-Syun Chen, Yung-Yu Chuang, Yen-Yu Lin
HIGHLIGHT: We address weakly supervised point cloud segmentation by proposing a new model, MIL-derived transformer, to mine additional supervisory signals.

698, TITLE: Fast Light-Weight Near-Field Photometric Stereo
http://openaccess.thecvf.com//content/CVPR2022/html/Lichy_Fast_Light-Weight_Near-Field_Photometric_Stereo_CVPR_2022_paper.html
AUTHORS: Daniel Lichy, Soumyadip Sengupta, David W. Jacobs
HIGHLIGHT: We introduce the first end-to-end learning-based solution to near-field Photometric Stereo (PS), where the light sources are close to the object of interest.

699, TITLE: BCOT: A Markerless High-Precision 3D Object Tracking Benchmark
http://openaccess.thecvf.com//content/CVPR2022/html/Li_BCOT_A_Markerless_High-Precision_3D_Object_Tracking_Benchmark_CVPR_2022_paper.html
AUTHORS: Jiachen Li, Bin Wang, Shiqiang Zhu, Xin Cao, Fan Zhong, Wenxuan Chen, Te Li, Jason Gu, Xueying Qin
HIGHLIGHT: In this paper, we present a multi-view approach to estimate the accurate 3D poses of real moving objects, and then use binocular data to construct a new benchmark for monocular textureless 3D object tracking.

700, TITLE: Omni-DETR: Omni-Supervised Object Detection With Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Omni-DETR_Omni-Supervised_Object_Detection_With_Transformers_CVPR_2022_paper.html
AUTHORS: Pei Wang, Zhaowei Cai, Hao Yang, Gurumurthy Swaminathan, Nuno Vasconcelos, Bernt Schiele, Stefano Soatto
HIGHLIGHT: We consider the problem of omni-supervised object detection, which can use unlabeled, fully labeled and weakly labeled annotations, such as image tags, counts, points, etc., for object detection.

701, TITLE: Uniform Subdivision of Omnidirectional Camera Space for Efficient Spherical Stereo Matching
http://openaccess.thecvf.com//content/CVPR2022/html/Kang_Uniform_Subdivision_of_Omnidirectional_Camera_Space_for_Efficient_Spherical_Stereo_CVPR_2022_paper.html
AUTHORS: Donghun Kang, Hyeonjoong Jang, Jungeon Lee, Chong-Min Kyung, Min H. Kim
HIGHLIGHT: In this paper, to address the distortion problem of omnidirectional images, we devise a novel subdivision scheme of a spherical geodesic grid.

702, TITLE: High-Resolution Image Synthesis With Latent Diffusion Models
http://openaccess.thecvf.com//content/CVPR2022/html/Rombach_High-Resolution_Image_Synthesis_With_Latent_Diffusion_Models_CVPR_2022_paper.html
AUTHORS: Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, Björn Ommer
HIGHLIGHT: By introducing cross-attention layers into the model architecture, we turn diffusion models into powerful and flexible generators for general conditioning inputs such as text or bounding boxes and high-resolution synthesis becomes possible in a convolutional manner.

703, TITLE: Improving Adversarially Robust Few-Shot Image Classification With Generalizable Representations
http://openaccess.thecvf.com//content/CVPR2022/html/Dong_Improving_Adversarially_Robust_Few-Shot_Image_Classification_With_Generalizable_Representations_CVPR_2022_paper.html
AUTHORS: Junhao Dong, Yuan Wang, Jian-Huang Lai, Xiaohua Xie
HIGHLIGHT: In this paper, we consider the FSIC problem in the case of adversarial examples.

704, TITLE: Transferable Sparse Adversarial Attack
http://openaccess.thecvf.com//content/CVPR2022/html/He_Transferable_Sparse_Adversarial_Attack_CVPR_2022_paper.html
AUTHORS: Ziwen He, Wei Wang, Jing Dong, Tieniu Tan
HIGHLIGHT: In this paper, we focus on sparse adversarial attack based on the l_0 norm constraint, which can succeed by only modifying a few pixels of an image.

705, TITLE: CREAM: Weakly Supervised Object Localization via Class RE-Activation Mapping
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_CREAM_Weakly_Supervised_Object_Localization_via_Class_RE-Activation_Mapping_CVPR_2022_paper.html
AUTHORS: Jilan Xu, Junlin Hou, Yuejie Zhang, Rui Feng, Rui-Wei Zhao, Tao Zhang, Xuequan Lu, Shang Gao
HIGHLIGHT: In this paper, we empirically prove that this problem is associated with the mixup of the activation values between less discriminative foreground regions and the background. To address it, we propose Class RE-Activation Mapping (CREAM), a novel clustering-based approach to boost the activation values of the integral object regions.

706, TITLE: Semi-Weakly-Supervised Learning of Complex Actions From Instructional Task Videos
http://openaccess.thecvf.com//content/CVPR2022/html/Shen_Semi-Weakly-Supervised_Learning_of_Complex_Actions_From_Instructional_Task_Videos_CVPR_2022_paper.html

AUTHORS: Yuhan Shen, Ehsan Elhamifar
HIGHLIGHT: We propose a general SWSL framework that can efficiently learn from both types of videos and can leverage any of the existing weakly-supervised action segmentation methods.

707, TITLE: APRIL: Finding the Achilles' Heel on Privacy for Vision Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Lu_APRIL_Finding_the_Achilles_Heel_on_Privacy_for_Vision_Transformers_CVPR_2022_paper.html
AUTHORS: Jiahao Lu, Xi Sheryl Zhang, Tianli Zhao, Xiangyu He, Jian Cheng
HIGHLIGHT: In this paper, we analyse the gradient leakage risk of self-attention based mechanism in both theoretical and practical manners.

708, TITLE: Text Spotting Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Text_Spotting_Transformers_CVPR_2022_paper.html
AUTHORS: Xiang Zhang, Yongwen Su, Subarna Tripathi, Zhuowen Tu
HIGHLIGHT: In this paper, we present TExt Spotting TRansformers (TESTR), a generic end-to-end text spotting framework using Transformers for text detection and recognition in the wild.

709, TITLE: Mip-NeRF 360: Unbounded Anti-Aliased Neural Radiance Fields
http://openaccess.thecvf.com//content/CVPR2022/html/Barron_Mip-NeRF_360_Unbounded_Anti-Aliased_Neural_Radiance_Fields_CVPR_2022_paper.html
AUTHORS: Jonathan T. Barron, Ben Mildenhall, Dor Verbin, Pratul P. Srinivasan, Peter Hedman
HIGHLIGHT: We present an extension of mip-NeRF (a NeRF variant that addresses sampling and aliasing) that uses a non-linear scene parameterization, online distillation, and a novel distortion-based regularizer to overcome the challenges presented by unbounded scenes.

710, TITLE: VALHALLA: Visual Hallucination for Machine Translation
http://openaccess.thecvf.com//content/CVPR2022/html/Li_VALHALLA_Visual_Hallucination_for_Machine_Translation_CVPR_2022_paper.html
AUTHORS: Yi Li, Rameswar Panda, Yoon Kim, Chun-Fu (Richard) Chen, Rogerio S. Feris, David Cox, Nuno Vasconcelos
HIGHLIGHT: In this paper, we introduce a visual hallucination framework, called VALHALLA, which requires only source sentences at inference time and instead uses hallucinated visual representations for multimodal machine translation.

711, TITLE: StyleSDF: High-Resolution 3D-Consistent Image and Geometry Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Or-El_StyleSDF_High-Resolution_3D-Consistent_Image_and_Geometry_Generation_CVPR_2022_paper.html
AUTHORS: Roy Or-El, Xuan Luo, Mengyi Shan, Eli Shechtman, Jeong Joon Park, Ira Kemelmacher-Shlizerman
HIGHLIGHT: We introduce a high resolution, 3D-consistent image and shape generation technique which we call StyleSDF.

712, TITLE: Incorporating Semi-Supervised and Positive-Unlabeled Learning for Boosting Full Reference Image Quality Assessment
http://openaccess.thecvf.com//content/CVPR2022/html/Cao_Incorporating_Semi-Supervised_and_Positive-Unlabeled_Learning_for_Boosting_Full_Reference_Image_CVPR_2022_paper.html
AUTHORS: Yue Cao, Zhaolin Wan, Dongwei Ren, Zifei Yan, Wangmeng Zuo
HIGHLIGHT: In this paper, we suggest to incorporate semi-supervised and positive-unlabeled (PU) learning for exploiting unlabeled data while mitigating the adverse effect of outliers.

713, TITLE: GLAMR: Global Occlusion-Aware Human Mesh Recovery With Dynamic Cameras
http://openaccess.thecvf.com//content/CVPR2022/html/Yuan_GLAMR_Global_Occlusion-Aware_Human_Mesh_Recovery_With_Dynamic_Cameras_CVPR_2022_paper.html
AUTHORS: Ye Yuan, Umar Iqbal, Pavlo Molchanov, Kris Kitani, Jan Kautz
HIGHLIGHT: We present an approach for 3D global human mesh recovery from monocular videos recorded with dynamic cameras.

714, TITLE: HINT: Hierarchical Neuron Concept Explainer
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_HINT_Hierarchical_Neuron_Concept_Explainer_CVPR_2022_paper.html
AUTHORS: Andong Wang, Wei-Ning Lee, Xiaojuan Qi
HIGHLIGHT: In this paper, we study hierarchical concepts inspired by the hierarchical cognition process of human beings.

715, TITLE: Capturing and Inferring Dense Full-Body Human-Scene Contact
http://openaccess.thecvf.com//content/CVPR2022/html/Huang_Capturing_and_Inferring_Dense_Full-Body_Human-Scene_Contact_CVPR_2022_paper.html
AUTHORS: Chun-Hao P. Huang, Hongwei Yi, Markus Höschle, Matvey Safroshkin, Tsvetelina Alexiadis, Senya Polikovsky, Daniel Scharstein, Michael J. Black
HIGHLIGHT: Our key insight is that regions in contact are always occluded so the network needs the ability to explore the whole image for evidence.

716, TITLE: Advancing High-Resolution Video-Language Representation With Large-Scale Video Transcriptions
http://openaccess.thecvf.com//content/CVPR2022/html/Xue_Advancing_High-Resolution_Video-Language_Representation_With_Large-Scale_Video_Transcriptions_CVPR_2022_paper.html

AUTHORS: Hongwei Xue, Tiankai Hang, Yanhong Zeng, Yuchong Sun, Bei Liu, Huan Yang, Jianlong Fu, Baining Guo
HIGHLIGHT: In this paper, we propose a novel High-resolution and Diversified VIdeo-LAnguage pre-training model (HD-VILA) for many visual tasks.

717, TITLE: Target-Aware Dual Adversarial Learning and a Multi-Scenario Multi-Modality Benchmark To Fuse Infrared and Visible for Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Target-Aware_Dual_Adversarial_Learning_and_a_Multi-Scenario_Multi-Modality_Benchmark_To_CVPR_2022_paper.html
AUTHORS: Jinyuan Liu, Xin Fan, Zhanbo Huang, Guanyao Wu, Risheng Liu, Wei Zhong, Zhongxuan Luo
HIGHLIGHT: This paper proposes a bilevel optimization formulation for the joint problem of fusion and detection, and then unrolls to a target-aware Dual Adversarial Learning (TarDAL) network for fusion and a commonly used detection network.

718, TITLE: En-Compactness: Self-Distillation Embedding & Contrastive Generation for Generalized Zero-Shot Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Kong_En-Compactness_Self-Distillation_Embedding__Contrastive_Generation_for_Generalized_Zero-Shot_Learning_CVPR_2022_paper.html
AUTHORS: Xia Kong, Zuodong Gao, Xiaofan Li, Ming Hong, Jun Liu, Chengjie Wang, Yuan Xie, Yanyun Qu
HIGHLIGHT: Instead, in this paper, we propose an Intra-Class Compactness Enhancement method (ICCE) for GZSL.

719, TITLE: Neural Face Identification in a 2D Wireframe Projection of a Manifold Object
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Neural_Face_Identification_in_a_2D_Wireframe_Projection_of_a_CVPR_2022_paper.html
AUTHORS: Kehan Wang, Jia Zheng, Zihan Zhou
HIGHLIGHT: In this paper, we approach the classical problem of face identification from a novel data-driven point of view.

720, TITLE: LC-FDNet: Learned Lossless Image Compression With Frequency Decomposition Network
http://openaccess.thecvf.com//content/CVPR2022/html/Rhee_LC-FDNet_Learned_Lossless_Image_Compression_With_Frequency_Decomposition_Network_CVPR_2022_paper.html
AUTHORS: Hochang Rhee, Yeong Il Jang, Seyun Kim, Nam Ik Cho
HIGHLIGHT: In this paper, we propose a new lossless image compression method that proceeds the encoding in a coarse-to-fine manner to separate and process low and high-frequency regions differently.

721, TITLE: Nonuniform-to-Uniform Quantization: Towards Accurate Quantization via Generalized Straight-Through Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Nonuniform-to-Uniform_Quantization_Towards_Accurate_Quantization_via_Generalized_Straight-Through_Estimation_CVPR_2022_paper.html
AUTHORS: Zechun Liu, Kwang-Ting Cheng, Dong Huang, Eric P. Xing, Zhiqiang Shen
HIGHLIGHT: In this study, we propose Nonuniform-to-Uniform Quantization (N2UQ), a method that can maintain the strong representation ability of nonuniform methods while being hardware-friendly and efficient as the uniform quantization for model inference.

722, TITLE: Deep Rectangling for Image Stitching: A Learning Baseline
http://openaccess.thecvf.com//content/CVPR2022/html/Nie_Deep_Rectangling_for_Image_Stitching_A_Learning_Baseline_CVPR_2022_paper.html
AUTHORS: Lang Nie, Chunyu Lin, Kang Liao, Shuaicheng Liu, Yao Zhao
HIGHLIGHT: However, these solutions only work for images with rich linear structures, leading to noticeable distortions for portraits and landscapes with non-linear objects. In this paper, we address these issues by proposing the first deep learning solution to image rectangling.

723, TITLE: PCL: Proxy-Based Contrastive Learning for Domain Generalization
http://openaccess.thecvf.com//content/CVPR2022/html/Yao_PCL_Proxy-Based_Contrastive_Learning_for_Domain_Generalization_CVPR_2022_paper.html
AUTHORS: Xufeng Yao, Yang Bai, Xinyun Zhang, Yuechen Zhang, Qi Sun, Ran Chen, Ruiyu Li, Bei Yu
HIGHLIGHT: We argue that aligning positive sample-to-sample pairs tends to hinder the model generalization due to the significant distribution gaps between different domains. To address this issue, we propose a novel proxy-based contrastive learning method, which replaces the original sample-to-sample relations with proxy-to-sample relations, significantly alleviating the positive alignment issue.

724, TITLE: SurfEmb: Dense and Continuous Correspondence Distributions for Object Pose Estimation With Learnt Surface Embeddings
http://openaccess.thecvf.com//content/CVPR2022/html/Haugaard_SurfEmb_Dense_and_Continuous_Correspondence_Distributions_for_Object_Pose_Estimation_CVPR_2022_paper.html
AUTHORS: Rasmus Laurvig Haugaard, Anders Glent Buch
HIGHLIGHT: We present an approach to learn dense, continuous 2D-3D correspondence distributions over the surface of objects from data with no prior knowledge of visual ambiguities like symmetry.

725, TITLE: Diverse Plausible 360-Degree Image Outpainting for Efficient 3DCG Background Creation
http://openaccess.thecvf.com//content/CVPR2022/html/Akimoto_Diverse_Plausible_360-Degree_Image_Outpainting_for_Efficient_3DCG_Background_Creation_CVPR_2022_paper.html
AUTHORS: Naofumi Akimoto, Yuhi Matsuo, Yoshimitsu Aoki

HIGHLIGHT: This paper proposes a completion method using a transformer for scene modeling and novel methods to improve the properties of a 360-degree image on the output image.

726, TITLE: Learning 3D Object Shape and Layout Without 3D Supervision
http://openaccess.thecvf.com//content/CVPR2022/html/Gkioxari_Learning_3D_Object_Shape_and_Layout_Without_3D_Supervision_CVPR_2022_paper.html
AUTHORS: Georgia Gkioxari, Nikhila Ravi, Justin Johnson
HIGHLIGHT: While there have been recent advances in predicting 3D shape and layout from a single image, most approaches rely on 3D ground truth for training which is expensive to collect at scale. We overcome these limitations and propose a method that learns to predict 3D shape and layout for objects without any ground truth shape or layout information: instead we rely on multi-view images with 2D supervision which can more easily be collected at scale.

727, TITLE: An Empirical Study of End-to-End Temporal Action Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_An_Empirical_Study_of_End-to-End_Temporal_Action_Detection_CVPR_2022_paper.html
AUTHORS: Xiaolong Liu, Song Bai, Xiang Bai
HIGHLIGHT: In this paper, we present an empirical study of end-to-end temporal action detection.

728, TITLE: SimVP: Simpler Yet Better Video Prediction
http://openaccess.thecvf.com//content/CVPR2022/html/Gao_SimVP_Simpler_Yet_Better_Video_Prediction_CVPR_2022_paper.html
AUTHORS: Zhangyang Gao, Cheng Tan, Lirong Wu, Stan Z. Li
HIGHLIGHT: This paper proposes SimVP, a simple video prediction model that is completely built upon CNN and trained by MSE loss in an end-to-end fashion.

729, TITLE: Object Localization Under Single Coarse Point Supervision
http://openaccess.thecvf.com//content/CVPR2022/html/Yu_Object_Localization_Under_Single_Coarse_Point_Supervision_CVPR_2022_paper.html
AUTHORS: Xuehui Yu, Pengfei Chen, Di Wu, Najmul Hassan, Guorong Li, Junchi Yan, Humphrey Shi, Qixiang Ye, Zhenjun Han
HIGHLIGHT: In this study, we propose a POL method using coarse point annotations, relaxing the supervision signals from accurate key points to freely spotted points.

730, TITLE: Unsupervised Learning of Accurate Siamese Tracking
http://openaccess.thecvf.com//content/CVPR2022/html/Shen_Unsupervised_Learning_of_Accurate_Siamese_Tracking_CVPR_2022_paper.html
AUTHORS: Qiuhong Shen, Lei Qiao, Jinyang Guo, Peixia Li, Xin Li, Bo Li, Weitao Feng, Weihao Gan, Wei Wu, Wanli Ouyang
HIGHLIGHT: We present a novel unsupervised tracking framework, in which we can learn temporal correspondence both on the classification branch and regression branch.

731, TITLE: Bayesian Nonparametric Submodular Video Partition for Robust Anomaly Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Sapkota_Bayesian_Nonparametric_Submodular_Video_Partition_for_Robust_Anomaly_Detection_CVPR_2022_paper.html
AUTHORS: Hitesh Sapkota, Qi Yu
HIGHLIGHT: We propose to conduct novel Bayesian non-parametric submodular video partition (BN-SVP) to significantly improve MIL model training that can offer a highly reliable solution for robust anomaly detection in practical settings that include outlier segments or multiple types of abnormal events.

732, TITLE: Brain-Supervised Image Editing
http://openaccess.thecvf.com//content/CVPR2022/html/Davis_Brain-Supervised_Image_Editing_CVPR_2022_paper.html
AUTHORS: Keith M. Davis III, Carlos de la Torre-Ortiz, Tuukka Ruotsalo
HIGHLIGHT: Here, we present a novel alternative: the utilization of brain responses as a supervision signal for learning semantic feature representations.

733, TITLE: 3D Shape Variational Autoencoder Latent Disentanglement via Mini-Batch Feature Swapping for Bodies and Faces
http://openaccess.thecvf.com//content/CVPR2022/html/Foti_3D_Shape_Variational_Autoencoder_Latent_Disentanglement_via_Mini-Batch_Feature_Swapping_CVPR_2022_paper.html
AUTHORS: Simone Foti, Bongjin Koo, Danail Stoyanov, Matthew J. Clarkson
HIGHLIGHT: In this paper, we propose an intuitive yet effective self-supervised approach to train a 3D shape variational autoencoder (VAE) which encourages a disentangled latent representation of identity features.

734, TITLE: Unified Transformer Tracker for Object Tracking
http://openaccess.thecvf.com//content/CVPR2022/html/Ma_Unified_Transformer_Tracker_for_Object_Tracking_CVPR_2022_paper.html
AUTHORS: Fan Ma, Mike Zheng Shou, Linchao Zhu, Haoqi Fan, Yilei Xu, Yi Yang, Zhicheng Yan
HIGHLIGHT: In this work, we present the Unified Transformer Tracker (UTT) to address tracking problems in different scenarios with one paradigm.

735, TITLE: Non-Parametric Depth Distribution Modelling Based Depth Inference for Multi-View Stereo

http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Non-Parametric_Depth_Distribution_Modelling_Based_Depth_Inference_for_Multi-View_Stereo_CVPR_2022_paper.html
AUTHORS: Jiayu Yang, Jose M. Alvarez, Miaomiao Liu
HIGHLIGHT: In contrast, we propose constructing the cost volume by non-parametric depth distribution modeling to handle pixels with unimodal and multi-modal distributions.

736, TITLE: Equalized Focal Loss for Dense Long-Tailed Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Equalized_Focal_Loss_for_Dense_Long-Tailed_Object_Detection_CVPR_2022_paper.html
AUTHORS: Bo Li, Yongqiang Yao, Jingru Tan, Gang Zhang, Fengwei Yu, Jianwei Lu, Ye Luo
HIGHLIGHT: However, in the long-tailed scenario, this line of work has not been explored so far. In this paper, we investigate whether one-stage detectors can perform well in this case.

737, TITLE: Generating High Fidelity Data From Low-Density Regions Using Diffusion Models
http://openaccess.thecvf.com//content/CVPR2022/html/Sehwag_Generating_High_Fidelity_Data_From_Low-Density_Regions_Using_Diffusion_Models_CVPR_2022_paper.html
AUTHORS: Vikash Sehwag, Caner Hazirbas, Albert Gordo, Firat Ozgenel, Cristian Canton
HIGHLIGHT: Our work focuses on addressing sample deficiency from low-density regions of data manifold in common image datasets.

738, TITLE: DeepDPM: Deep Clustering With an Unknown Number of Clusters
http://openaccess.thecvf.com//content/CVPR2022/html/Ronen_DeepDPM_Deep_Clustering_With_an_Unknown_Number_of_Clusters_CVPR_2022_paper.html
AUTHORS: Meitar Ronen, Shahaf E. Finder, Oren Freifeld
HIGHLIGHT: When K is unknown, however, using model-selection criteria to choose its optimal value might become computationally expensive, especially in DL as the training process would have to be repeated numerous times. In this work, we bridge this gap by introducing an effective deep-clustering method that does not require knowing the value of K as it infers it during the learning.

739, TITLE: Spiking Transformers for Event-Based Single Object Tracking
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Spiking_Transformers_for_Event-Based_Single_Object_Tracking_CVPR_2022_paper.html
AUTHORS: Jiqing Zhang, Bo Dong, Haiwei Zhang, Jianchuan Ding, Felix Heide, Baocai Yin, Xin Yang
HIGHLIGHT: In this work, we propose a spiking transformer network, STNet, for single object tracking.

740, TITLE: FocalClick: Towards Practical Interactive Image Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_FocalClick_Towards_Practical_Interactive_Image_Segmentation_CVPR_2022_paper.html
AUTHORS: Xi Chen, Zhiyan Zhao, Yilei Zhang, Manni Duan, Donglian Qi, Hengshuang Zhao
HIGHLIGHT: Although explored by many previous works, there is still a gap between academic approaches and industrial needs: first, existing models are not efficient enough to work on low power devices; second, they perform poorly when used to refine preexisting masks as they could not avoid destroying the correct part. FocalClick solves both issues at once by predicting and updating the mask in localized areas.

741, TITLE: ISDNet: Integrating Shallow and Deep Networks for Efficient Ultra-High Resolution Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Guo_ISDNet_Integrating_Shallow_and_Deep_Networks_for_Efficient_Ultra-High_Resolution_CVPR_2022_paper.html
AUTHORS: Shaohua Guo, Liang Liu, Zhenye Gan, Yabiao Wang, Wuhao Zhang, Chengjie Wang, Guannan Jiang, Wei Zhang, Ran Yi, Lizhuang Ma, Ke Xu
HIGHLIGHT: In this paper, we propose ISDNet, a novel ultra-high resolution segmentation framework that integrates the shallow and deep networks in a new manner, which significantly accelerates the inference speed while achieving accurate segmentation.

742, TITLE: Unsupervised Domain Adaptation for Nighttime Aerial Tracking
http://openaccess.thecvf.com//content/CVPR2022/html/Ye_Unsupervised_Domain_Adaptation_for_Nighttime_Aerial_Tracking_CVPR_2022_paper.html
AUTHORS: Junjie Ye, Changhong Fu, Guangze Zheng, Danda Pani Paudel, Guang Chen
HIGHLIGHT: This work instead develops a novel unsupervised domain adaptation framework for nighttime aerial tracking (named UDAT).

743, TITLE: Balanced Multimodal Learning via On-the-Fly Gradient Modulation
http://openaccess.thecvf.com//content/CVPR2022/html/Peng_Balanced_Multimodal_Learning_via_On-the-Fly_Gradient_Modulation_CVPR_2022_paper.html
AUTHORS: Xiaokang Peng, Yake Wei, Andong Deng, Dong Wang, Di Hu
HIGHLIGHT: Specifically, in this paper we point out that existing audio-visual discriminative models, in which uniform objective is designed for all modalities, could remain under-optimized uni-modal representations, caused by another dominated modality in some scenarios, e.g., sound in blowing wind event, vision in drawing picture event, etc. To alleviate this optimization imbalance, we propose on-the-fly gradient modulation to adaptively control the optimization of each modality, via monitoring the discrepancy of their contribution towards the learning objective.

744, TITLE: RestoreFormer: High-Quality Blind Face Restoration From Undegraded Key-Value Pairs
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_RestoreFormer_High-
Quality_Blind_Face_Restoration_From_Undegraded_Key-Value_Pairs_CVPR_2022_paper.html
AUTHORS: Zhouxia Wang, Jiawei Zhang, Runjian Chen, Wenping Wang, Ping Luo
HIGHLIGHT: As face image contains abundant contextual information, we propose a method, RestoreFormer, which explores fully-spatial attentions to model contextual information and surpasses existing works that use local convolutions.

745, TITLE: Understanding Uncertainty Maps in Vision With Statistical Testing
http://openaccess.thecvf.com//content/CVPR2022/html/Nazarovs_Understanding_Uncertainty_Maps_in_Vision_With_Statistical_Testing_CVPR_2022_paper.html
AUTHORS: Jurijs Nazarovs, Zhichun Huang, Songwong Tasneeyapant, Rudrasis Chakraborty, Vikas Singh
HIGHLIGHT: In this paper, specifically for uncertainties defined on images, we show how revisiting results from Random Field theory (RFT) when paired with DNN tools (to get around computational hurdles) leads to efficient frameworks that can provide a hypothesis test capabilities, not otherwise available, for uncertainty maps from models used in many vision tasks.

746, TITLE: CAFE: Learning To Condense Dataset by Aligning Features
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_CAFE_Learning_To_Condense_Dataset_by_Aligning_Features_CVPR_2022_paper.html
AUTHORS: Kai Wang, Bo Zhao, Xiangyu Peng, Zheng Zhu, Shuo Yang, Shuo Wang, Guan Huang, Hakan Bilen, Xinchao Wang, Yang You
HIGHLIGHT: In this paper, we propose a novel scheme to Condense dataset by Aligning FEatures (CAFE), which explicitly attempts to preserve the real-feature distribution as well as the discriminant power of the resulting synthetic set, lending itself to strong generalization capability to various architectures.

747, TITLE: Causality Inspired Representation Learning for Domain Generalization
http://openaccess.thecvf.com//content/CVPR2022/html/Lv_Causality_Inspired_Representation_Learning_for_Domain_Generalization_CVPR_2022_paper.html
AUTHORS: Fangrui Lv, Jian Liang, Shuang Li, Bin Zang, Chi Harold Liu, Ziteng Wang, Di Liu
HIGHLIGHT: When the dependence changes with the target distribution, the statistic models may fail to generalize. In this regard, we introduce a general structural causal model to formalize the DG problem.

748, TITLE: Mask-Guided Spectral-Wise Transformer for Efficient Hyperspectral Image Reconstruction
http://openaccess.thecvf.com//content/CVPR2022/html/Cai_Mask-Guided_Spectral-
Wise_Transformer_for_Efficient_Hyperspectral_Image_Reconstruction_CVPR_2022_paper.html
AUTHORS: Yuanhao Cai, Jing Lin, Xiaowan Hu, Haoqian Wang, Xin Yuan, Yulun Zhang, Radu Timofte, Luc Van Gool
HIGHLIGHT: In this paper, we propose a novel framework, Mask-guided Spectral-wise Transformer (MST), for HSI reconstruction.

749, TITLE: A Variational Bayesian Method for Similarity Learning in Non-Rigid Image Registration
http://openaccess.thecvf.com//content/CVPR2022/html/Grzech_A_Variational_Bayesian_Method_for_Similarity_Learning_in_Non-Rigid_Image_CVPR_2022_paper.html
AUTHORS: Daniel Grzech, Mohammad Farid Azampour, Ben Glocker, Julia Schnabel, Nassir Navab, Bernhard Kainz, Loïc Le Folgoc
HIGHLIGHT: We propose a novel variational Bayesian formulation for diffeomorphic non-rigid registration of medical images, which learns in an unsupervised way a data-specific similarity metric.

750, TITLE: Not Just Selection, but Exploration: Online Class-Incremental Continual Learning via Dual View Consistency
http://openaccess.thecvf.com//content/CVPR2022/html/Gu_Not_Just_Selection_but_Exploration_Online_Class-
Incremental_Continual_Learning_via_CVPR_2022_paper.html
AUTHORS: Yanan Gu, Xu Yang, Kun Wei, Cheng Deng
HIGHLIGHT: In this paper, we propose a novel yet effective framework for online class-incremental continual learning, which considers not only the selection of stored samples, but also the full exploration of the data stream.

751, TITLE: PPDL: Predicate Probability Distribution Based Loss for Unbiased Scene Graph Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Li_PPDL_Predicate_Probability_Distribution_Based_Loss_for_Unbiased_Scene_Graph_CVPR_2022_paper.html
AUTHORS: Wei Li, Haiwei Zhang, Qijie Bai, Guoqing Zhao, Ning Jiang, Xiaojie Yuan
HIGHLIGHT: In this paper, we propose a Predicate Probability Distribution based Loss (PPDL) to train the biased SGG models and obtain unbiased Scene Graphs ultimately.

752, TITLE: Block-NeRF: Scalable Large Scene Neural View Synthesis
http://openaccess.thecvf.com//content/CVPR2022/html/Tancik_Block-
NeRF_Scalable_Large_Scene_Neural_View_Synthesis_CVPR_2022_paper.html
AUTHORS: Matthew Tancik, Vincent Casser, Xinchen Yan, Sabeek Pradhan, Ben Mildenhall, Pratul P. Srinivasan, Jonathan T. Barron, Henrik Kretzschmar
HIGHLIGHT: We present Block-NeRF, a variant of Neural Radiance Fields that can represent large-scale environments.

753, TITLE: Coupling Vision and Proprioception for Navigation of Legged Robots
http://openaccess.thecvf.com//content/CVPR2022/html/Fu_Coupling_Vision_and_Proprioception_for_Navigation_of_Legged_Robots_CVPR_2022_paper.html

AUTHORS: Zipeng Fu, Ashish Kumar, Ananye Agarwal, Haozhi Qi, Jitendra Malik, Deepak Pathak
HIGHLIGHT: We exploit the complementary strengths of vision and proprioception to develop a point-goal navigation system for legged robots, called VP-Nav.

754, TITLE: Fine-Grained Predicates Learning for Scene Graph Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Lyu_Fine-Grained_Predicates_Learning_for_Scene_Graph_Generation_CVPR_2022_paper.html
AUTHORS: Xinyu Lyu, Lianli Gao, Yuyu Guo, Zhou Zhao, Hao Huang, Heng Tao Shen, Jingkuan Song
HIGHLIGHT: While general SGG models are prone to predict head predicates and existing re-balancing strategies prefer tail categories, none of them can appropriately handle these hard-to-distinguish predicates. To tackle this issue, inspired by fine-grained image classification, which focuses on differentiating among hard-to-distinguish object classes, we propose a method named Fine-Grained Predicates Learning (FGPL) which aims at differentiating among hard-to-distinguish predicates for Scene Graph Generation task.

755, TITLE: Generalized Few-Shot Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Tian_Generalized_Few-Shot_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Zhuotao Tian, Xin Lai, Li Jiang, Shu Liu, Michelle Shu, Hengshuang Zhao, Jiaya Jia
HIGHLIGHT: In this paper, we introduce a new benchmark, called Generalized Few-Shot Semantic Segmentation (GFS-Seg), to analyze the generalization ability of simultaneously segmenting the novel categories with very few examples and the base categories with sufficient examples.

756, TITLE: Exploiting Rigidity Constraints for LiDAR Scene Flow Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Dong_Exploiting_Rigidity_Constraints_for_LiDAR_Scene_Flow_Estimation_CVPR_2022_paper.html
AUTHORS: Guanting Dong, Yueyi Zhang, Hanlin Li, Xiaoyan Sun, Zhiwei Xiong
HIGHLIGHT: In this paper, we propose a novel optimization method based on a recurrent neural network to predict LiDAR scene flow in a weakly supervised manner.

757, TITLE: Neural Head Avatars From Monocular RGB Videos
http://openaccess.thecvf.com//content/CVPR2022/html/Grassal_Neural_Head_Avatars_From_Monocular_RGB_Videos_CVPR_2022_paper.html
AUTHORS: Philip-William Grassal, Malte Prinzler, Titus Leistner, Carsten Rother, Matthias Nießner, Justus Thies
HIGHLIGHT: We present Neural Head Avatars, a novel neural representation that explicitly models the surface geometry and appearance of an animatable human avatar that can be used for teleconferencing in AR/VR or other applications in the movie or games industry that rely on a digital human.

758, TITLE: B-Cos Networks: Alignment Is All We Need for Interpretability
http://openaccess.thecvf.com//content/CVPR2022/html/Bohle_B-Cos_Networks_Alignment_Is_All_We_Need_for_Interpretability_CVPR_2022_paper.html
AUTHORS: Moritz Böhle, Mario Fritz, Bernt Schiele
HIGHLIGHT: We present a new direction for increasing the interpretability of deep neural networks (DNNs) by promoting weight-input alignment during training.

759, TITLE: EMOCA: Emotion Driven Monocular Face Capture and Animation
http://openaccess.thecvf.com//content/CVPR2022/html/Danecek_EMOCA_Emotion_Driven_Monocular_Face_Capture_and_Animation_CVPR_2022_paper.html
AUTHORS: Radek Daněček, Michael J. Black, Timo Bolkart
HIGHLIGHT: The result is facial geometries that do not match the emotional content of the input image. We address this with EMOCA (EMOtion Capture and Animation), by introducing a novel deep perceptual emotion consistency loss during training, which helps ensure that the reconstructed 3D expression matches the expression depicted in the input image.

760, TITLE: Burst Image Restoration and Enhancement
http://openaccess.thecvf.com//content/CVPR2022/html/Dudhane_Burst_Image_Restoration_and_Enhancement_CVPR_2022_paper.html
AUTHORS: Akshay Dudhane, Syed Waqas Zamir, Salman Khan, Fahad Shahbaz Khan, Ming-Hsuan Yang
HIGHLIGHT: Our central idea is to create a set of pseudo-burst features that combine complimentary information from all the input burst frames to seamlessly exchange information.

761, TITLE: What Makes Transfer Learning Work for Medical Images: Feature Reuse & Other Factors
http://openaccess.thecvf.com//content/CVPR2022/html/Matsoukas_What_Makes_Transfer_Learning_Work_for_Medical_Images_Feature_Reuse_CVPR_2022_paper.html
AUTHORS: Christos Matsoukas, Johan Fredin Haslum, Moein Sorkhei, Magnus Söderberg, Kevin Smith
HIGHLIGHT: Through a series of experiments on several medical image benchmark datasets, we explore the relationship between transfer learning, data size, the capacity and inductive bias of the model, as well as the distance between the source and target domain. Our findings suggest that transfer learning is beneficial in most cases, and we characterize the important role feature reuse plays in its success.

762, TITLE: Towards Diverse and Natural Scene-Aware 3D Human Motion Synthesis

http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Towards_Diverse_and_Natural_Scene-Aware_3D_Human_Motion_Synthesis_CVPR_2022_paper.html
AUTHORS: Jingbo Wang, Yu Rong, Jingyuan Liu, Sijie Yan, Dahua Lin, Bo Dai
HIGHLIGHT: In this paper, we focus on the problem of synthesizing diverse scene-aware human motions under the guidance of target action sequences.

763, TITLE: Quarantine: Sparsity Can Uncover the Trojan Attack Trigger for Free
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Quarantine_Sparsity_Can_Uncover_the_Trojan_Attack_Trigger_for_Free_CVPR_2022_paper.html
AUTHORS: Tianlong Chen, Zhenyu Zhang, Yihua Zhang, Shiyu Chang, Sijia Liu, Zhangyang Wang
HIGHLIGHT: Our crucial observation is that the Trojan features are significantly more stable to network pruning than benign features. Leveraging that, we propose a novel Trojan network detection regime: first locating a "winning Trojan lottery ticket" which preserves nearly full Trojan information yet only chance-level performance on clean inputs; then recovering the trigger embedded in this already isolated subnetwork.

764, TITLE: Audio-Visual Speech Codecs: Rethinking Audio-Visual Speech Enhancement by Re-Synthesis
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Audio-Visual_Speech_Codecs_Rethinking_Audio-Visual_Speech_Enhancement_by_Re-Synthesis_CVPR_2022_paper.html
AUTHORS: Karren Yang, Dejan MarkoviÄ‡, Steven Krenn, Vasu Agrawal, Alexander Richard
HIGHLIGHT: In this paper, we propose a novel audio-visual speech enhancement framework for high-fidelity telecommunications in AR/VR.

765, TITLE: Localized Adversarial Domain Generalization
http://openaccess.thecvf.com//content/CVPR2022/html/Zhu_Localized_Adversarial_Domain_Generalization_CVPR_2022_paper.html
AUTHORS: Wei Zhu, Le Lu, Jing Xiao, Mei Han, Jiebo Luo, Adam P. Harrison
HIGHLIGHT: Adversarial domain generalization is a popular approach to DG, but conventional approaches (1) struggle to sufficiently align features so that local neighborhoods are mixed across domains; and (2) can suffer from feature space over collapse which can threaten generalization performance. To address these limitations, we propose localized adversarial domain generalization with space compactness maintenance (LADG) which constitutes two major contributions. First, we propose an adversarial localized classifier as the domain discriminator, along with a principled primary branch.

766, TITLE: X-Trans2Cap: Cross-Modal Knowledge Transfer Using Transformer for 3D Dense Captioning
http://openaccess.thecvf.com//content/CVPR2022/html/Yuan_X-Trans2Cap_Cross-Modal_Knowledge_Transfer_Using_Transformer_for_3D_Dense_Captioning_CVPR_2022_paper.html
AUTHORS: Zhihao Yuan, Xu Yan, Yinghong Liao, Yao Guo, Guanbin Li, Shuguang Cui, Zhen Li
HIGHLIGHT: In this study, we investigate a cross-modal knowledge transfer using Transformer for 3D dense captioning, X-Trans2Cap, to effectively boost the performance of single-modal 3D caption through knowledge distillation using a teacher-student framework.

767, TITLE: How Much Does Input Data Type Impact Final Face Model Accuracy?
http://openaccess.thecvf.com//content/CVPR2022/html/Luo_How_Much_Does_Input_Data_Type_Impact_Final_Face_Model_CVPR_2022_paper.html
AUTHORS: Jiahao Luo, Fahim Hasan Khan, Issei Mori, Akila de Silva, Eric Sandoval Ruezga, Minghao Liu, Alex Pang, James Davis
HIGHLIGHT: If a particular application domain requires accuracy X, which kinds of input data are suitable? Does the input data need to be 3D, or will 2D data suffice? This paper takes a step toward answering these questions using synthetic data.

768, TITLE: Image-to-Lidar Self-Supervised Distillation for Autonomous Driving Data
http://openaccess.thecvf.com//content/CVPR2022/html/Sautier_Image-to-Lidar_Self-Supervised_Distillation_for_Autonomous_Driving_Data_CVPR_2022_paper.html
AUTHORS: Corentin Sautier, Gilles Puy, Spyros Gidaris, Alexandre Boulch, Andrei Bursuc, Renaud Marlet
HIGHLIGHT: Yet annotating 3D Lidar data for these tasks is tedious and costly. In this context, we propose a self-supervised pre-training method for 3D perception models that is tailored to autonomous driving data.

769, TITLE: HumanNeRF: Free-Viewpoint Rendering of Moving People From Monocular Video
http://openaccess.thecvf.com//content/CVPR2022/html/Weng_HumanNeRF_Free-Viewpoint_Rendering_of_Moving_People_From_Monocular_Video_CVPR_2022_paper.html
AUTHORS: Chung-Yi Weng, Brian Curless, Pratul P. Srinivasan, Jonathan T. Barron, Ira Kemelmacher-Shlizerman
HIGHLIGHT: We introduce a free-viewpoint rendering method -- HumanNeRF -- that works on a given monocular video of a human performing complex body motions, e.g. a video from YouTube.

770, TITLE: PoseKernelLifter: Metric Lifting of 3D Human Pose Using Sound
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_PoseKernelLifter_Metric_Lifting_of_3D_Human_Pose_Using_Sound_CVPR_2022_paper.html
AUTHORS: Zhijian Yang, Xiaoran Fan, Volkan Isler, Hyun Soo Park
HIGHLIGHT: In this paper, we show that audio signals recorded along with an image, provide complementary information to reconstruct the metric 3D pose of the person.

771, TITLE: Which Images To Label for Few-Shot Medical Landmark Detection?

82

http://openaccess.thecvf.com//content/CVPR2022/html/Quan_Which_Images_To_Label_for_Few-Shot_Medical_Landmark_Detection_CVPR_2022_paper.html
AUTHORS: Quan Quan, Qingsong Yao, Jun Li, S. Kevin Zhou
HIGHLIGHT: We herein propose a novel Sample Choosing Policy (SCP) to select "the most worthy" images as the templates, in the context of medical landmark detection.

772, TITLE: Why Discard if You Can Recycle?: A Recycling Max Pooling Module for 3D Point Cloud Analysis
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Why_Discard_if_You_Can_Recycle_A_Recycling_Max_Pooling_CVPR_2022_paper.html
AUTHORS: Jiajing Chen, Burak Kakillioglu, Huantao Ren, Senem Velipasalar
HIGHLIGHT: We first show that this traditional approach causes only a fraction of 3D points contribute to the permutation invariant features, and discards the rest of the points. In order to address this issue and improve the performance of any baseline 3D point classification or segmentation model, we propose a new module, referred to as the Recycling MaxPooling (RMP) module, to recycle and utilize the features of some of the discarded points.

773, TITLE: Explaining Deep Convolutional Neural Networks via Latent Visual-Semantic Filter Attention
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Explaining_Deep_Convolutional_Neural_Networks_via_Latent_Visual-Semantic_Filter_Attention_CVPR_2022_paper.html
AUTHORS: Yu Yang, Seungbae Kim, Jungseock Joo
HIGHLIGHT: However, generating semantic explanations about the learned representation is challenging without direct supervision to produce such explanations. We propose a general framework, Latent Visual Semantic Explainer (LaViSE), to teach any existing convolutional neural network to generate text descriptions about its own latent representations at the filter level.

774, TITLE: AlignQ: Alignment Quantization With ADMM-Based Correlation Preservation
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_AlignQ_Alignment_Quantization_With_ADMM-Based_Correlation_Preservation_CVPR_2022_paper.html
AUTHORS: Ting-An Chen, De-Nian Yang, Ming-Syan Chen
HIGHLIGHT: However, existing approaches ignored the distribution difference between training and testing data, thereby inducing a large quantization error in inference. To address this issue, we propose a new quantization scheme, Alignment Quantization with ADMM-based Correlation Preservation (AlignQ), which exploits the cumulative distribution function (CDF) to align the data to be i.i.d. (independently and identically distributed) for quantization error minimization.

775, TITLE: Self-Distillation From the Last Mini-Batch for Consistency Regularization
http://openaccess.thecvf.com//content/CVPR2022/html/Shen_Self-Distillation_From_the_Last_Mini-Batch_for_Consistency_Regularization_CVPR_2022_paper.html
AUTHORS: Yiqing Shen, Liwu Xu, Yuzhe Yang, Yaqian Li, Yandong Guo
HIGHLIGHT: However, they either require extra network architecture modification, or are difficult to parallelize. To cope with these challenges, we propose an efficient and reliable self-distillation framework, named Self-Distillation from Last Mini-Batch (DLB).

776, TITLE: Interactive Multi-Class Tiny-Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Lee_Interactive_Multi-Class_Tiny-Object_Detection_CVPR_2022_paper.html
AUTHORS: Chunggi Lee, Seonwook Park, Heon Song, Jeongun Ryu, Sanghoon Kim, Haejoon Kim, Sérgio Pereira, Donggeun Yoo
HIGHLIGHT: Such imagery typically contains objects from various categories, yet the multi-class interactive annotation setting for the detection task has thus far been unexplored. To address these needs, we propose a novel interactive annotation method for multiple instances of tiny objects from multiple classes, based on a few point-based user inputs.

777, TITLE: Learning From Pixel-Level Noisy Label: A New Perspective for Light Field Saliency Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Feng_Learning_From_Pixel-Level_Noisy_Label_A_New_Perspective_for_Light_CVPR_2022_paper.html
AUTHORS: Mingtao Feng, Kendong Liu, Liang Zhang, Hongshan Yu, Yaonan Wang, Ajmal Mian
HIGHLIGHT: In this paper, we propose to learn light field saliency from pixel-level noisy labels obtained from unsupervised hand crafted featured-based saliency methods.

778, TITLE: UBoCo: Unsupervised Boundary Contrastive Learning for Generic Event Boundary Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Kang_UBoCo_Unsupervised_Boundary_Contrastive_Learning_for_Generic_Event_Boundary_Detection_CVPR_2022_paper.html
AUTHORS: Hyolim Kang, Jinwoo Kim, Taehyun Kim, Seon Joo Kim
HIGHLIGHT: In this paper, we propose a novel framework for unsupervised/supervised GEBD, by using the Temporal Self-similarity Matrix (TSM) as the video representation.

779, TITLE: Multi-View Depth Estimation by Fusing Single-View Depth Probability With Multi-View Geometry
http://openaccess.thecvf.com//content/CVPR2022/html/Bae_Multi-View_Depth_Estimation_by_Fusing_Single-View_Depth_Probability_With_Multi-View_CVPR_2022_paper.html
AUTHORS: Gwangbin Bae, Ignas Budvytis, Roberto Cipolla
HIGHLIGHT: For such failure modes, single-view depth estimation methods are often more reliable. To this end, we propose MaGNet, a novel framework for fusing single-view depth probability with multi-view geometry, to improve the accuracy, robustness and efficiency of multi-view depth estimation.

780, TITLE: Learning To Collaborate in Decentralized Learning of Personalized Models
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Learning_To_Collaborate_in_Decentralized_Learning_of_Personalized_Models_CVPR_2022_paper.html
AUTHORS: Shuangtong Li, Tianyi Zhou, Xinmei Tian, Dacheng Tao
HIGHLIGHT: In this paper, we dynamically update the mixing-weights to improve the personalized model for each node's task and meanwhile learn a sparse topology to reduce communication costs.

781, TITLE: CLIP-NeRF: Text-and-Image Driven Manipulation of Neural Radiance Fields
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_CLIP-NeRF_Text-and-Image_Driven_Manipulation_of_Neural_Radiance_Fields_CVPR_2022_paper.html
AUTHORS: Can Wang, Menglei Chai, Mingming He, Dongdong Chen, Jing Liao
HIGHLIGHT: We present CLIP-NeRF, a multi-modal 3D object manipulation method for neural radiance fields (NeRF).

782, TITLE: ART-Point: Improving Rotation Robustness of Point Cloud Classifiers via Adversarial Rotation
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_ART-Point_Improving_Rotation_Robustness_of_Point_Cloud_Classifiers_via_Adversarial_CVPR_2022_paper.html
AUTHORS: Ruibin Wang, Yibo Yang, Dacheng Tao
HIGHLIGHT: In this study, for the first time, we show that the rotation robustness of point cloud classifiers can also be acquired via adversarial training with better performance on both rotated and clean datasets.

783, TITLE: Ref-NeRF: Structured View-Dependent Appearance for Neural Radiance Fields
http://openaccess.thecvf.com//content/CVPR2022/html/Verbin_Ref-NeRF_Structured_View-Dependent_Appearance_for_Neural_Radiance_Fields_CVPR_2022_paper.html
AUTHORS: Dor Verbin, Peter Hedman, Ben Mildenhall, Todd Zickler, Jonathan T. Barron, Pratul P. Srinivasan
HIGHLIGHT: While NeRF-based techniques excel at representing fine geometric structures with smoothly varying view-dependent appearance, they often fail to accurately capture and reproduce the appearance of glossy surfaces. We address this limitation by introducing Ref-NeRF, which replaces NeRF's parameterization of view-dependent outgoing radiance with a representation of reflected radiance and structures this function using a collection of spatially-varying scene properties.

784, TITLE: 360-Attack: Distortion-Aware Perturbations From Perspective-Views
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_360-Attack_Distortion-Aware_Perturbations_From_Perspective-Views_CVPR_2022_paper.html
AUTHORS: Yunjian Zhang, Yanwei Liu, Jinxia Liu, Jingbo Miao, Antonios Argyriou, Liming Wang, Zhen Xu
HIGHLIGHT: In this paper, we propose an adversarial attack targeting spherical images, called 360-attactk, that transfers adversarial perturbations from perspective-view (PV) images to a final adversarial spherical image.

785, TITLE: Targeted Supervised Contrastive Learning for Long-Tailed Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Targeted_Supervised_Contrastive_Learning_for_Long-Tailed_Recognition_CVPR_2022_paper.html
AUTHORS: Tianhong Li, Peng Cao, Yuan Yuan, Lijie Fan, Yuzhe Yang, Rogerio S. Feris, Piotr Indyk, Dina Katabi
HIGHLIGHT: In this paper, we show that while supervised contrastive learning can help improve performance, past baselines suffer from poor uniformity brought in by imbalanced data distribution.

786, TITLE: Both Style and Fog Matter: Cumulative Domain Adaptation for Semantic Foggy Scene Understanding
http://openaccess.thecvf.com//content/CVPR2022/html/Ma_Both_Style_and_Fog_Matter_Cumulative_Domain_Adaptation_for_Semantic_CVPR_2022_paper.html
AUTHORS: Xianzheng Ma, Zhixiang Wang, Yacheng Zhan, Yinqiang Zheng, Zheng Wang, Dengxin Dai, Chia-Wen Lin
HIGHLIGHT: Thus, we propose a new pipeline to cumulatively adapt style, fog and the dual-factor (style and fog).

787, TITLE: Ev-TTA: Test-Time Adaptation for Event-Based Object Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Kim_Ev-TTA_Test-Time_Adaptation_for_Event-Based_Object_Recognition_CVPR_2022_paper.html
AUTHORS: Junho Kim, Inwoo Hwang, Young Min Kim
HIGHLIGHT: We introduce Ev-TTA, a simple, effective test-time adaptation algorithm for event-based object recognition.

788, TITLE: Balanced Contrastive Learning for Long-Tailed Visual Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Zhu_Balanced_Contrastive_Learning_for_Long-Tailed_Visual_Recognition_CVPR_2022_paper.html
AUTHORS: Jianggang Zhu, Zheng Wang, Jingjing Chen, Yi-Ping Phoebe Chen, Yu-Gang Jiang
HIGHLIGHT: In this paper, we focus on representation learning for imbalanced data.

789, TITLE: Slimmable Domain Adaptation
http://openaccess.thecvf.com//content/CVPR2022/html/Meng_Slimmable_Domain_Adaptation_CVPR_2022_paper.html
AUTHORS: Rang Meng, Weijie Chen, Shicai Yang, Jie Song, Luojun Lin, Di Xie, Shiliang Pu, Xinchao Wang, Mingli Song, Yueting Zhuang
HIGHLIGHT: In this paper, we introduce a simple framework, Slimmable Domain Adaptation, to improve cross-domain generalization with a weight-sharing model bank, from which models of different capacities can be sampled to accommodate different accuracy-efficiency trade-offs.

790, TITLE: Bandits for Structure Perturbation-Based Black-Box Attacks To Graph Neural Networks With Theoretical Guarantees
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Bandits_for_Structure_Perturbation-Based_Black-Box_Attacks_To_Graph_Neural_Networks_CVPR_2022_paper.html
AUTHORS: Binghui Wang, Youqi Li, Pan Zhou
HIGHLIGHT: Existing attacks to GNNs are either under the less practical threat model where the attacker is assumed to access the GNN model parameters, or under the practical black-box threat model but consider perturbing node features that are shown to be not enough effective. In this paper, we aim to bridge this gap and consider black-box attacks to GNNs with structure perturbation as well as with theoretical guarantees.

791, TITLE: NODEO: A Neural Ordinary Differential Equation Based Optimization Framework for Deformable Image Registration
http://openaccess.thecvf.com//content/CVPR2022/html/Wu_NODEO_A_Neural_Ordinary_Differential_Equation_Based_Optimization_Framework_for_CVPR_2022_paper.html
AUTHORS: Yifan Wu, Tom Z. Jiahao, Jiancong Wang, Paul A. Yushkevich, M. Ani Hsieh, James C. Gee
HIGHLIGHT: In this paper, we present a novel, generic, and accurate diffeomorphic image registration framework that utilizes neural ordinary differential equations (NODEs).

792, TITLE: DIP: Deep Inverse Patchmatch for High-Resolution Optical Flow
http://openaccess.thecvf.com//content/CVPR2022/html/Zheng_DIP_Deep_Inverse_Patchmatch_for_High-Resolution_Optical_Flow_CVPR_2022_paper.html
AUTHORS: Zihua Zheng, Ni Nie, Zhi Ling, Pengfei Xiong, Jiangyu Liu, Hao Wang, Jiankun Li
HIGHLIGHT: In this paper, we propose a novel Patchmatch-based framework to work on high-resolution optical flow estimation.

793, TITLE: Few-Shot Object Detection With Fully Cross-Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Han_Few-Shot_Object_Detection_With_Fully_Cross-Transformer_CVPR_2022_paper.html
AUTHORS: Guangxing Han, Jiawei Ma, Shiyuan Huang, Long Chen, Shih-Fu Chang
HIGHLIGHT: Inspired by the recent work on vision transformers and vision-language transformers, we propose a novel Fully Cross-Transformer based model (FCT) for FSOD by incorporating cross-transformer into both the feature backbone and detection head.

794, TITLE: Pyramid Architecture for Multi-Scale Processing in Point Cloud Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Nie_Pyramid_Architecture_for_Multi-Scale_Processing_in_Point_Cloud_Segmentation_CVPR_2022_paper.html
AUTHORS: Dong Nie, Rui Lan, Ling Wang, Xiaofeng Ren
HIGHLIGHT: In this work, we investigate how to efficiently and effectively integrate features at varying scales and varying stages in a point cloud segmentation network.

795, TITLE: Decoupling Makes Weakly Supervised Local Feature Better
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Decoupling_Makes_Weakly_Supervised_Local_Feature_Better_CVPR_2022_paper.html
AUTHORS: Kunhong Li, Longguang Wang, Li Liu, Qing Ran, Kai Xu, Yulan Guo
HIGHLIGHT: In this paper, we propose a decoupled training describe-then-detect pipeline tailored for weakly supervised local feature learning.

796, TITLE: Cross-Architecture Self-Supervised Video Representation Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Guo_Cross-Architecture_Self-Supervised_Video_Representation_Learning_CVPR_2022_paper.html
AUTHORS: Sheng Guo, Zihua Xiong, Yujie Zhong, Limin Wang, Xiaobo Guo, Bing Han, Weilin Huang
HIGHLIGHT: In this paper, we present a new cross-architecture contrastive learning (CACL) framework for self-supervised video representation learning.

797, TITLE: High-Resolution Image Harmonization via Collaborative Dual Transformations
http://openaccess.thecvf.com//content/CVPR2022/html/Cong_High-Resolution_Image_Harmonization_via_Collaborative_Dual_Transformations_CVPR_2022_paper.html
AUTHORS: Wenyan Cong, Xinhao Tao, Li Niu, Jing Liang, Xuesong Gao, Qihao Sun, Liqing Zhang
HIGHLIGHT: In this work, we propose a high-resolution image harmonization network with Collaborative Dual Transformation (CDTNet) to combine pixel-to-pixel transformation and RGB-to-RGB transformation coherently in an end-to-end network.

798, TITLE: Homography Loss for Monocular 3D Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Gu_Homography_Loss_for_Monocular_3D_Object_Detection_CVPR_2022_paper.html
AUTHORS: Jiaqi Gu, Bojian Wu, Lubin Fan, Jianqiang Huang, Shen Cao, Zhiyu Xiang, Xian-Sheng Hua
HIGHLIGHT: In this paper, we propose a novel method that takes all the objects into consideration and explores their mutual relationships to help better estimate the 3D boxes.

799, TITLE: A Unified Model for Line Projections in Catadioptric Cameras With Rotationally Symmetric Mirrors

http://openaccess.thecvf.com//content/CVPR2022/html/Miraldo_A_Unified_Model_for_Line_Projections_in_Catadioptric_Cameras_With_CVPR_2022_paper.html
AUTHORS: Pedro Miraldo, José Pedro Iglesias
HIGHLIGHT: We start by taking some general point reflection assumptions and derive a line reflection constraint. This constraint is then used to define a line projection into the image.

800, TITLE: Dynamic Sparse R-CNN
http://openaccess.thecvf.com//content/CVPR2022/html/Hong_Dynamic_Sparse_R-CNN_CVPR_2022_paper.html
AUTHORS: Qinghang Hong, Fengming Liu, Dong Li, Ji Liu, Lu Tian, Yi Shan
HIGHLIGHT: In this work, we propose to improve Sparse R-CNN with two dynamic designs.

801, TITLE: MM-TTA: Multi-Modal Test-Time Adaptation for 3D Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Shin_MM-TTA_Multi-Modal_Test-Time_Adaptation_for_3D_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Inkyu Shin, Yi-Hsuan Tsai, Bingbing Zhuang, Samuel Schulter, Buyu Liu, Sparsh Garg, In So Kweon, Kuk-Jin Yoon
HIGHLIGHT: In this paper, we propose and explore a new multi-modal extension of test-time adaptation for 3D semantic segmentation.

802, TITLE: Stable Long-Term Recurrent Video Super-Resolution
http://openaccess.thecvf.com//content/CVPR2022/html/Chiche_Stable_Long-Term_Recurrent_Video_Super-Resolution_CVPR_2022_paper.html
AUTHORS: Benjamin Naoto Chiche, Arnaud Woiselle, Joana Frontera-Pons, Jean-Luc Starck
HIGHLIGHT: In this work, we expose instabilities of existing recurrent VSR networks on long sequences with low motion.

803, TITLE: Dual-Generator Face Reenactment
http://openaccess.thecvf.com//content/CVPR2022/html/Hsu_Dual-Generator_Face_Reenactment_CVPR_2022_paper.html
AUTHORS: Gee-Sern Hsu, Chun-Hung Tsai, Hung-Yi Wu
HIGHLIGHT: We propose the Dual-Generator (DG) network for large-pose face reenactment.

804, TITLE: Towards Bidirectional Arbitrary Image Rescaling: Joint Optimization and Cycle Idempotence
http://openaccess.thecvf.com//content/CVPR2022/html/Pan_Towards_Bidirectional_Arbitrary_Image_Rescaling_Joint_Optimization_and_Cycle_Idempotence_CVPR_2022_paper.html
AUTHORS: Zhihong Pan, Baopu Li, Dongliang He, Mingde Yao, Wenhao Wu, Tianwei Lin, Xin Li, Errui Ding
HIGHLIGHT: Our proposed method is the first to treat arbitrary rescaling, both upscaling and downscaling, as one unified process.

805, TITLE: Self-Supervised Neural Articulated Shape and Appearance Models
http://openaccess.thecvf.com//content/CVPR2022/html/Wei_Self-Supervised_Neural_Articulated_Shape_and_Appearance_Models_CVPR_2022_paper.html
AUTHORS: Fangyin Wei, Rohan Chabra, Lingni Ma, Christoph Lassner, Michael Zollhöfer, Szymon Rusinkiewicz, Chris Sweeney, Richard Newcombe, Mira Slavcheva
HIGHLIGHT: We propose a novel approach for learning a representation of the geometry, appearance, and motion of a class of articulated objects given only a set of color images as input.

806, TITLE: A Hybrid Quantum-Classical Algorithm for Robust Fitting
http://openaccess.thecvf.com//content/CVPR2022/html/Doan_A_Hybrid_Quantum-Classical_Algorithm_for_Robust_Fitting_CVPR_2022_paper.html
AUTHORS: Anh-Dzung Doan, Michele Sasdelli, David Suter, Tat-Jun Chin
HIGHLIGHT: In this paper, we propose a hybrid quantum-classical algorithm for robust fitting.

807, TITLE: Topology Preserving Local Road Network Estimation From Single Onboard Camera Image
http://openaccess.thecvf.com//content/CVPR2022/html/Can_Topology_Preserving_Local_Road_Network_Estimation_From_Single_Onboard_Camera_CVPR_2022_paper.html
AUTHORS: Yigit Baran Can, Alexander Liniger, Danda Pani Paudel, Luc Van Gool
HIGHLIGHT: This paper aims at extracting the local road network topology, directly in the bird's-eye-view (BEV), all in a complex urban setting.

808, TITLE: Eigenlanes: Data-Driven Lane Descriptors for Structurally Diverse Lanes
http://openaccess.thecvf.com//content/CVPR2022/html/Jin_Eigenlanes_Data-Driven_Lane_Descriptors_for_Structurally_Diverse_Lanes_CVPR_2022_paper.html
AUTHORS: Dongkwon Jin, Wonhui Park, Seong-Gyun Jeong, Heeyeon Kwon, Chang-Su Kim
HIGHLIGHT: A novel algorithm to detect road lanes in the eigenlane space is proposed in this paper.

809, TITLE: Human Instance Matting via Mutual Guidance and Multi-Instance Refinement
http://openaccess.thecvf.com//content/CVPR2022/html/Sun_Human_Instance_Matting_via_Mutual_Guidance_and_Multi-Instance_Refinement_CVPR_2022_paper.html
AUTHORS: Yanan Sun, Chi-Keung Tang, Yu-Wing Tai
HIGHLIGHT: This paper introduces a new matting task called human instance matting (HIM), which requires the pertinent model to automatically predict a precise alpha matte for each human instance.

810, TITLE: TCTrack: Temporal Contexts for Aerial Tracking
http://openaccess.thecvf.com//content/CVPR2022/html/Cao_TCTrack_Temporal_Contexts_for_Aerial_Tracking_CVPR_2022_paper.html
AUTHORS: Ziang Cao, Ziyuan Huang, Liang Pan, Shiwei Zhang, Ziwei Liu, Changhong Fu
HIGHLIGHT: In this work, we present TCTrack, a comprehensive framework to fully exploit temporal contexts for aerial tracking.

811, TITLE: SpaceEdit: Learning a Unified Editing Space for Open-Domain Image Color Editing
http://openaccess.thecvf.com//content/CVPR2022/html/Shi_SpaceEdit_Learning_a_Unified_Editing_Space_for_Open-Domain_Image_Color_CVPR_2022_paper.html
AUTHORS: Jing Shi, Ning Xu, Haitian Zheng, Alex Smith, Jiebo Luo, Chenliang Xu
HIGHLIGHT: Recently, large pretrained models (e.g., BERT, StyleGAN, CLIP) show great knowledge transfer and generalization capability on various downstream tasks within their domains. Inspired by these efforts, in this paper we propose a unified model for open-domain image editing focusing on color and tone adjustment of open-domain images while keeping their original content and structure.

812, TITLE: GAN-Supervised Dense Visual Alignment
http://openaccess.thecvf.com//content/CVPR2022/html/Peebles_GAN-Supervised_Dense_Visual_Alignment_CVPR_2022_paper.html
AUTHORS: William Peebles, Jun-Yan Zhu, Richard Zhang, Antonio Torralba, Alexei A. Efros, Eli Shechtman
HIGHLIGHT: We propose GAN-Supervised Learning, a framework for learning discriminative models and their GAN-generated training data jointly end-to-end.

813, TITLE: SwinTextSpotter: Scene Text Spotting via Better Synergy Between Text Detection and Text Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Huang_SwinTextSpotter_Scene_Text_Spotting_via_Better_Synergy_Between_Text_Detection_CVPR_2022_paper.html
AUTHORS: Mingxin Huang, Yuliang Liu, Zhenghao Peng, Chongyu Liu, Dahua Lin, Shenggao Zhu, Nicholas Yuan, Kai Ding, Lianwen Jin
HIGHLIGHT: In this paper, we propose a new end-to-end scene text spotting framework termed SwinTextSpotter.

814, TITLE: Multi-Level Feature Learning for Contrastive Multi-View Clustering
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_Multi-Level_Feature_Learning_for_Contrastive_Multi-View_Clustering_CVPR_2022_paper.html
AUTHORS: Jie Xu, Huayi Tang, Yazhou Ren, Liang Peng, Xiaofeng Zhu, Lifang He
HIGHLIGHT: In this paper, we propose a new framework of multi-level feature learning for contrastive multi-view clustering to address the aforementioned issue.

815, TITLE: RendNet: Unified 2D/3D Recognizer With Latent Space Rendering
http://openaccess.thecvf.com//content/CVPR2022/html/Shi_RendNet_Unified_2D3D_Recognizer_With_Latent_Space_Rendering_CVPR_2022_paper.html
AUTHORS: Ruoxi Shi, Xinyang Jiang, Caihua Shan, Yansen Wang, Dongsheng Li
HIGHLIGHT: As a result, we propose RenderNet, a unified architecture for recognition on both 2D and 3D scenarios, which considers both VG/RG representations and exploits their interaction by incorporating the VG-to-RG rasterization process.

816, TITLE: iPLAN: Interactive and Procedural Layout Planning
http://openaccess.thecvf.com//content/CVPR2022/html/He_iPLAN_Interactive_and_Procedural_Layout_Planning_CVPR_2022_paper.html
AUTHORS: Feixiang He, Yanlong Huang, He Wang
HIGHLIGHT: However, the capability of involving humans into the loop has been largely ignored in existing methods which are mostly end-to-end approaches. To this end, we propose a new human-in-the-loop generative model, iPLAN, which is capable of automatically generating layouts, but also interacting with designers throughout the whole procedure, enabling humans and AI to co-evolve a sketchy idea gradually into the final design.

817, TITLE: Video Frame Interpolation With Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Lu_Video_Frame_Interpolation_With_Transformer_CVPR_2022_paper.html
AUTHORS: Liying Lu, Ruizheng Wu, Huaijia Lin, Jiangbo Lu, Jiaya Jia
HIGHLIGHT: Existing methods built upon convolutional networks generally face challenges of handling large motion due to the locality of convolution operations. To overcome this limitation, we introduce a novel framework, which takes advantage of Transformer to model long-range pixel correlation among video frames.

818, TITLE: GIFS: Neural Implicit Function for General Shape Representation
http://openaccess.thecvf.com//content/CVPR2022/html/Ye_GIFS_Neural_Implicit_Function_for_General_Shape_Representation_CVPR_2022_paper.html
AUTHORS: Jianglong Ye, Yuntao Chen, Naiyan Wang, Xiaolong Wang
HIGHLIGHT: In this work, we propose a novel method to represent general shapes including non-watertight shapes and shapes with multi-layer surfaces.

819, TITLE: Deblur-NeRF: Neural Radiance Fields From Blurry Images

http://openaccess.thecvf.com//content/CVPR2022/html/Ma_Deblur-NeRF_Neural_Radiance_Fields_From_Blurry_Images_CVPR_2022_paper.html
AUTHORS: Li Ma, Xiaoyu Li, Jing Liao, Qi Zhang, Xuan Wang, Jue Wang, Pedro V. Sander
HIGHLIGHT: However, image blurriness caused by defocus or motion, which often occurs when capturing scenes in the wild, significantly degrades its reconstruction quality. To address this problem, We propose Deblur-NeRF, the first method that can recover a sharp NeRF from blurry input.

820, TITLE: Egocentric Prediction of Action Target in 3D
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Egocentric_Prediction_of_Action_Target_in_3D_CVPR_2022_paper.html
AUTHORS: Yiming Li, Ziang Cao, Andrew Liang, Benjamin Liang, Luoyao Chen, Hang Zhao, Chen Feng
HIGHLIGHT: It is important in fields like human-robot collaboration, but has not yet received enough attention from vision and learning communities. To stimulate more research on this challenging egocentric vision task, we propose a large multimodality dataset of more than 1 million frames of RGB-D and IMU streams, and provide evaluation metrics based on our high-quality 2D and 3D labels from semi-automatic annotation.

821, TITLE: TemporalUV: Capturing Loose Clothing With Temporally Coherent UV Coordinates
http://openaccess.thecvf.com//content/CVPR2022/html/Xie_TemporalUV_Capturing_Loose_Clothing_With_Temporally_Coherent_UV_Coordinates_CVPR_2022_paper.html
AUTHORS: You Xie, Huiqi Mao, Angela Yao, Nils Thuerey
HIGHLIGHT: We propose a novel approach to generate temporally coherent UV coordinates for loose clothing.

822, TITLE: Whose Track Is It Anyway? Improving Robustness to Tracking Errors With Affinity-Based Trajectory Prediction
http://openaccess.thecvf.com//content/CVPR2022/html/Weng_Whose_Track_Is_It_Anyway_Improving_Robustness_to_Tracking_Errors_CVPR_2022_paper.html
AUTHORS: Xinshuo Weng, Boris Ivanovic, Kris Kitani, Marco Pavone
HIGHLIGHT: This is typically caused by the propagation of errors from tracking to prediction, such as noisy tracks, fragments, and identity switches. To alleviate this propagation of errors, we propose a new prediction paradigm that uses detections and their affinity matrices across frames as inputs, removing the need for error-prone data association during tracking.

823, TITLE: DoubleField: Bridging the Neural Surface and Radiance Fields for High-Fidelity Human Reconstruction and Rendering
http://openaccess.thecvf.com//content/CVPR2022/html/Shao_DoubleField_Bridging_the_Neural_Surface_and_Radiance_Fields_for_High-Fidelity_CVPR_2022_paper.html
AUTHORS: Ruizhi Shao, Hongwen Zhang, He Zhang, Mingjia Chen, Yan-Pei Cao, Tao Yu, Yebin Liu
HIGHLIGHT: We introduce DoubleField, a novel framework combining the merits of both surface field and radiance field for high-fidelity human reconstruction and rendering.

824, TITLE: Towards Real-World Navigation With Deep Differentiable Planners
http://openaccess.thecvf.com//content/CVPR2022/html/Ishida_Towards_Real-World_Navigation_With_Deep_Differentiable_Planners_CVPR_2022_paper.html
AUTHORS: Shu Ishida, Joã*o F. Henriques
HIGHLIGHT: We train embodied neural networks to plan and navigate unseen complex 3D environments, emphasising real-world deployment.

825, TITLE: An Iterative Quantum Approach for Transformation Estimation From Point Sets
http://openaccess.thecvf.com//content/CVPR2022/html/Meli_An_Iterative_Quantum_Approach_for_Transformation_Estimation_From_Point_Sets_CVPR_2022_paper.html
AUTHORS: Natacha Kuete Meli, Florian Mannel, Jan Lellmann
HIGHLIGHT: We propose an iterative method for estimating rigid transformations from point sets using adiabatic quantum computation.

826, TITLE: Video K-Net: A Simple, Strong, and Unified Baseline for Video Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Video_K-Net_A_Simple_Strong_and_Unified_Baseline_for_Video_CVPR_2022_paper.html
AUTHORS: Xiangtai Li, Wenwei Zhang, Jiangmiao Pang, Kai Chen, Guangliang Cheng, Yunhai Tong, Chen Change Loy
HIGHLIGHT: This paper presents Video K-Net, a simple, strong, and unified framework for fully end-to-end video panoptic segmentation.

827, TITLE: UnweaveNet: Unweaving Activity Stories
http://openaccess.thecvf.com//content/CVPR2022/html/Price_UnweaveNet_Unweaving_Activity_Stories_CVPR_2022_paper.html
AUTHORS: Will Price, Carl Vondrick, Dima Damen
HIGHLIGHT: We propose and showcase the efficacy of pretraining UnweaveNet in a self-supervised manner.

828, TITLE: Balanced MSE for Imbalanced Visual Regression
http://openaccess.thecvf.com//content/CVPR2022/html/Ren_Balanced_MSE_for_Imbalanced_Visual_Regression_CVPR_2022_paper.html
AUTHORS: Jiawei Ren, Mingyuan Zhang, Cunjun Yu, Ziwei Liu
HIGHLIGHT: In this work, we identify that the widely used Mean Square Error (MSE) loss function can be ineffective in imbalanced regression.

829, TITLE:	Local Learning Matters: Rethinking Data Heterogeneity in Federated Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Mendieta_Local_Learning_Matters_Rethinking_Data_Heterogeneity_in_Federated_Learning_CVPR_2022_paper.html
AUTHORS:	Matias Mendieta, Taojiannan Yang, Pu Wang, Minwoo Lee, Zhengming Ding, Chen Chen
HIGHLIGHT:	Instead, we consider rethinking solutions to data heterogeneity in FL with a focus on local learning generality rather than proximal restriction.

830, TITLE:	PhysFormer: Facial Video-Based Physiological Measurement With Temporal Difference Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Yu_PhysFormer_Facial_Video-Based_Physiological_Measurement_With_Temporal_Difference_Transformer_CVPR_2022_paper.html
AUTHORS:	Zitong Yu, Yuming Shen, Jingang Shi, Hengshuang Zhao, Philip H.S. Torr, Guoying Zhao
HIGHLIGHT:	In this paper, we propose the PhysFormer, an end-to-end video transformer based architecture, to adaptively aggregate both local and global spatio-temporal features for rPPG representation enhancement.

831, TITLE:	Dimension Embeddings for Monocular 3D Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Dimension_Embeddings_for_Monocular_3D_Object_Detection_CVPR_2022_paper.html
AUTHORS:	Yunpeng Zhang, Wenzhao Zheng, Zheng Zhu, Guan Huang, Dalong Du, Jie Zhou, Jiwen Lu
HIGHLIGHT:	In this paper, we propose a general method to learn appropriate embeddings for dimension estimation in monocular 3D object detection.

832, TITLE:	Look Closer To Supervise Better: One-Shot Font Generation via Component-Based Discriminator
http://openaccess.thecvf.com//content/CVPR2022/html/Kong_Look_Closer_To_Supervise_Better_One-Shot_Font_Generation_via_Component-Based_CVPR_2022_paper.html
AUTHORS:	Yuxin Kong, Canjie Luo, Weihong Ma, Qiyuan Zhu, Shenggao Zhu, Nicholas Yuan, Lianwen Jin
HIGHLIGHT:	We investigate the drawbacks of previous studies and find that a coarse-grained discriminator is insufficient for supervising a font generator. To this end, we propose a novel Component-Aware Module (CAM), which supervises the generator to decouple content and style at a more fine-grained level, i.e., the component level.

833, TITLE:	NeRFReN: Neural Radiance Fields With Reflections
http://openaccess.thecvf.com//content/CVPR2022/html/Guo_NeRFReN_Neural_Radiance_Fields_With_Reflections_CVPR_2022_paper.html
AUTHORS:	Yuan-Chen Guo, Di Kang, Linchao Bao, Yu He, Song-Hai Zhang
HIGHLIGHT:	Specifically, we propose to split a scene into transmitted and reflected components, and model the two components with separate neural radiance fields.

834, TITLE:	Blind Image Super-Resolution With Elaborate Degradation Modeling on Noise and Kernel
http://openaccess.thecvf.com//content/CVPR2022/html/Yue_Blind_Image_Super-Resolution_With_Elaborate_Degradation_Modeling_on_Noise_and_CVPR_2022_paper.html
AUTHORS:	Zongsheng Yue, Qian Zhao, Jianwen Xie, Lei Zhang, Deyu Meng, Kwan-Yee K. Wong
HIGHLIGHT:	To address the above issues, this paper proposes a model-based blind SISR method under the probabilistic framework, which elaborately models image degradation from the perspectives of noise and blur kernel.

835, TITLE:	Finding Good Configurations of Planar Primitives in Unorganized Point Clouds
http://openaccess.thecvf.com//content/CVPR2022/html/Yu_Finding_Good_Configurations_of_Planar_Primitives_in_Unorganized_Point_Clouds_CVPR_2022_paper.html
AUTHORS:	Mulin Yu, Florent Lafarge
HIGHLIGHT:	We present an algorithm for detecting planar primitives from unorganized 3D point clouds.

836, TITLE:	PhyIR: Physics-Based Inverse Rendering for Panoramic Indoor Images
http://openaccess.thecvf.com//content/CVPR2022/html/Li_PhyIR_Physics-Based_Inverse_Rendering_for_Panoramic_Indoor_Images_CVPR_2022_paper.html
AUTHORS:	Zhen Li, Lingli Wang, Xiang Huang, Cihui Pan, Jiaqi Yang
HIGHLIGHT:	In this paper, we present PhyIR, a neural inverse rendering method with a more completed SVBRDF representation and a physics-based in-network rendering layer, which can handle complex material and incorporate physical constraints by re-rendering realistic and detailed specular reflectance.

837, TITLE:	SCS-Co: Self-Consistent Style Contrastive Learning for Image Harmonization
http://openaccess.thecvf.com//content/CVPR2022/html/Hang_SCS-Co_Self-Consistent_Style_Contrastive_Learning_for_Image_Harmonization_CVPR_2022_paper.html
AUTHORS:	Yucheng Hang, Bin Xia, Wenming Yang, Qingmin Liao
HIGHLIGHT:	Moreover, recent region-aware adaptive instance normalization achieves great success but only considers the global background feature distribution, making the aligned foreground feature distribution biased. To address these issues, we propose a self-consistent style contrastive learning scheme (SCS-Co).

838, TITLE:	Beyond Fixation: Dynamic Window Visual Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Ren_Beyond_Fixation_Dynamic_Window_Visual_Transformer_CVPR_2022_paper.html
AUTHORS:	Pengzhen Ren, Changlin Li, Guangrun Wang, Yun Xiao, Qing Du, Xiaodan Liang, Xiaojun Chang

HIGHLIGHT: In this paper, we propose a novel method, named Dynamic Window Vision Transformer (DW-ViT).

839, TITLE: Progressive End-to-End Object Detection in Crowded Scenes
http://openaccess.thecvf.com//content/CVPR2022/html/Zheng_Progressive_End-to-End_Object_Detection_in_Crowded_Scenes_CVPR_2022_paper.html
AUTHORS: Anlin Zheng, Yuang Zhang, Xiangyu Zhang, Xiaojuan Qi, Jian Sun
HIGHLIGHT: In this paper, we propose a new query-based detection framework for crowd detection.

840, TITLE: FMCNet: Feature-Level Modality Compensation for Visible-Infrared Person Re-Identification
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_FMCNet_Feature-Level_Modality_Compensation_for_Visible-Infrared_Person_Re-Identification_CVPR_2022_paper.html
AUTHORS: Qiang Zhang, Changzhou Lai, Jianan Liu, Nianchang Huang, Jungong Han
HIGHLIGHT: Alternatively, we present a novel Feature-level Modality Compensation Network (FMCNet) for VIReID in this paper, which aims to compensate the missing modality-specific information in the feature level rather than in the image level, i.e., directly generating those missing modality-specific features of one modality from existing modality-shared features of the other modality.

841, TITLE: Improving GAN Equilibrium by Raising Spatial Awareness
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Improving_GAN_Equilibrium_by_Raising_Spatial_Awareness_CVPR_2022_paper.html
AUTHORS: Jianyuan Wang, Ceyuan Yang, Yinghao Xu, Yujun Shen, Hongdong Li, Bolei Zhou
HIGHLIGHT: To alleviate the issue of D dominating the competition in GANs, we aim to raise the spatial awareness of G. Randomly sampled multi-level heatmaps are encoded into the intermediate layers of G as an inductive bias.

842, TITLE: Neural Convolutional Surfaces
http://openaccess.thecvf.com//content/CVPR2022/html/Morreale_Neural_Convolutional_Surfaces_CVPR_2022_paper.html
AUTHORS: Luca Morreale, Noam Aigerman, Paul Guerrero, Vladimir G. Kim, Niloy J. Mitra
HIGHLIGHT: This work is concerned with representation of shapes while disentangling fine, local and possibly repeating geometry, from global, coarse structures.

843, TITLE: HyperSegNAS: Bridging One-Shot Neural Architecture Search With 3D Medical Image Segmentation Using HyperNet
http://openaccess.thecvf.com//content/CVPR2022/html/Peng_HyperSegNAS_Bridging_One-Shot_Neural_Architecture_Search_With_3D_Medical_Image_Segmentation_CVPR_2022_paper.html
AUTHORS: Cheng Peng, Andriy Myronenko, Ali Hatamizadeh, Vishwesh Nath, Md Mahfuzur Rahman Siddiquee, Yufan He, Daguang Xu, Rama Chellappa, Dong Yang
HIGHLIGHT: To enable one-shot NAS for medical image segmentation, our method, named HyperSegNAS, introduces a HyperNet to assist super-net training by incorporating architecture topology information.

844, TITLE: A Comprehensive Study of Image Classification Model Sensitivity to Foregrounds, Backgrounds, and Visual Attributes
http://openaccess.thecvf.com//content/CVPR2022/html/Moayeri_A_Comprehensive_Study_of_Image_Classification_Model_Sensitivity_to_Foregrounds_CVPR_2022_paper.html
AUTHORS: Mazda Moayeri, Phillip Pope, Yogesh Balaji, Soheil Feizi
HIGHLIGHT: We find that, somewhat surprisingly, in ResNets, adversarial training makes models more sensitive to the background compared to foreground than standard training.

845, TITLE: ConDor: Self-Supervised Canonicalization of 3D Pose for Partial Shapes
http://openaccess.thecvf.com//content/CVPR2022/html/Sajnani_ConDor_Self-Supervised_Canonicalization_of_3D_Pose_for_Partial_Shapes_CVPR_2022_paper.html
AUTHORS: Rahul Sajnani, Adrien Poulenard, Jivitesh Jain, Radhika Dua, Leonidas J. Guibas, Srinath Sridhar
HIGHLIGHT: ConDor is a self-supervised method that learns to Canonicalize the 3D orientation and position for full and partial 3D point clouds.

846, TITLE: Source-Free Domain Adaptation via Distribution Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Ding_Source-Free_Domain_Adaptation_via_Distribution_Estimation_CVPR_2022_paper.html
AUTHORS: Ning Ding, Yixing Xu, Yehui Tang, Chao Xu, Yunhe Wang, Dacheng Tao
HIGHLIGHT: In this work, we propose a novel framework called SFDA-DE to address SFDA task via source Distribution Estimation.

847, TITLE: Robust Combination of Distributed Gradients Under Adversarial Perturbations
http://openaccess.thecvf.com//content/CVPR2022/html/Kim_Robust_Combination_of_Distributed_Gradients_Under_Adversarial_Perturbations_CVPR_2022_paper.html
AUTHORS: Kwang In Kim
HIGHLIGHT: We propose a new server-side learning algorithm that robustly combines gradients.

848, TITLE: Exploring Endogenous Shift for Cross-Domain Detection: A Large-Scale Benchmark and Perturbation Suppression Network

http://openaccess.thecvf.com//content/CVPR2022/html/Tao_Exploring_Endogenous_Shift_for_Cross-Domain_Detection_A_Large-Scale_Benchmark_and_CVPR_2022_paper.html
AUTHORS: Renshuai Tao, Hainan Li, Tianbo Wang, Yanlu Wei, Yifu Ding, Bowei Jin, Hongping Zhi, Xianglong Liu, Aishan Liu
HIGHLIGHT: To support this study, we contribute the first Endogenous Domain Shift (EDS) benchmark, X-ray security inspection, where the endogenous shifts among the domains are mainly caused by different X-ray machine types with different hardware parameters, wear degrees, etc.

849, TITLE: VisCUIT: Visual Auditor for Bias in CNN Image Classifier
http://openaccess.thecvf.com//content/CVPR2022/html/Lee_VisCUIT_Visual_Auditor_for_Bias_in_CNN_Image_Classifier_CVPR_2022_paper.html
AUTHORS: Seongmin Lee, Zijie J. Wang, Judy Hoffman, Duen Horng (Polo) Chau
HIGHLIGHT: We present VisCUIT, an interactive visualization system that reveals how and why a CNN classifier is biased.

850, TITLE: Automatic Synthesis of Diverse Weak Supervision Sources for Behavior Analysis
http://openaccess.thecvf.com//content/CVPR2022/html/Tseng_Automatic_Synthesis_of_Diverse_Weak_Supervision_Sources_for_Behavior_Analysis_CVPR_2022_paper.html
AUTHORS: Albert Tseng, Jennifer J. Sun, Yisong Yue
HIGHLIGHT: To reduce expert effort, we present AutoSWAP: a framework for automatically synthesizing data-efficient task-level LFs.

851, TITLE: Transferability Estimation Using Bhattacharyya Class Separability
http://openaccess.thecvf.com//content/CVPR2022/html/Pandy_Transferability_Estimation_Using_Bhattacharyya_Class_Separability_CVPR_2022_paper.html
AUTHORS: Michal Pándy, Andrea Agostinelli, Jasper Uijlings, Vittorio Ferrari, Thomas Mensink
HIGHLIGHT: In this work, we propose Gaussian Bhattacharyya Coefficient (GBC), a novel method for quantifying transferability between a source model and a target dataset.

852, TITLE: DirecFormer: A Directed Attention in Transformer Approach to Robust Action Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Truong_DirecFormer_A_Directed_Attention_in_Transformer_Approach_to_Robust_Action_CVPR_2022_paper.html
AUTHORS: Thanh-Dat Truong, Quoc-Huy Bui, Chi Nhan Duong, Han-Seok Seo, Son Lam Phung, Xin Li, Khoa Luu
HIGHLIGHT: This work presents a novelend-to-end Transformer-based Directed Attention (Direc-Former) framework1for robust action recognition.

853, TITLE: Hierarchical Self-Supervised Representation Learning for Movie Understanding
http://openaccess.thecvf.com//content/CVPR2022/html/Xiao_Hierarchical_Self-Supervised_Representation_Learning_for_Movie_Understanding_CVPR_2022_paper.html
AUTHORS: Fanyi Xiao, Kaustav Kundu, Joseph Tighe, Davide Modolo
HIGHLIGHT: In contrast, in this paper we focus on self-supervised video learning for movie understanding and propose a novel hierarchical self-supervised pretraining strategy that separately pretrains each level of our hierarchical movie understanding model.

854, TITLE: Robust Egocentric Photo-Realistic Facial Expression Transfer for Virtual Reality
http://openaccess.thecvf.com//content/CVPR2022/html/Jourabloo_Robust_Egocentric_Photo-Realistic_Facial_Expression_Transfer_for_Virtual_Reality_CVPR_2022_paper.html
AUTHORS: Amin Jourabloo, Fernando De la Torre, Jason Saragih, Shih-En Wei, Stephen Lombardi, Te-Li Wang, Danielle Belko, Autumn Trimble, Hernan Badino
HIGHLIGHT: Our key contribution to improve robustness and generalization, is that our method implicitly decouples, in an unsupervised manner, the facial expression from nuisance factors (e.g., headset, environment, facial appearance).

855, TITLE: Does Robustness on ImageNet Transfer to Downstream Tasks?
http://openaccess.thecvf.com//content/CVPR2022/html/Yamada_Does_Robustness_on_ImageNet_Transfer_to_Downstream_Tasks_CVPR_2022_paper.html
AUTHORS: Yutaro Yamada, Mayu Otani
HIGHLIGHT: For object detection and semantic segmentation, we find that a vanilla Swin Transformer, a variant of Vision Transformer tailored for dense prediction tasks, transfers robustness better than Convolutional Neural Networks that are trained to be robust to the corrupted version of ImageNet.

856, TITLE: Propagation Regularizer for Semi-Supervised Learning With Extremely Scarce Labeled Samples
http://openaccess.thecvf.com//content/CVPR2022/html/Kim_Propagation_Regularizer_for_Semi-Supervised_Learning_With_Extremely_Scarce_Labeled_Samples_CVPR_2022_paper.html
AUTHORS: Noo-ri Kim, Jee-Hyong Lee
HIGHLIGHT: In this paper, we focus on SSL environment with extremely scarce labeled samples, only 1 or 2 labeled samples per class, where most of existing methods fail to learn.

857, TITLE: Bailando: 3D Dance Generation by Actor-Critic GPT With Choreographic Memory
http://openaccess.thecvf.com//content/CVPR2022/html/Siyao_Bailando_3D_Dance_Generation_by_Actor-Critic_GPT_With_Choreographic_Memory_CVPR_2022_paper.html
AUTHORS: Li Siyao, Weijiang Yu, Tianpei Gu, Chunze Lin, Quan Wang, Chen Qian, Chen Change Loy, Ziwei Liu

HIGHLIGHT: In addition, the generated dance sequence also needs to maintain temporal coherency with different music genres. To tackle these challenges, we propose a novel music-to-dance framework, Bailando, with two powerful components: 1) a choreographic memory that learns to summarize meaningful dancing units from 3D pose sequence to a quantized codebook, 2) an actor-critic Generative Pre-trained Transformer (GPT) that composes these units to a fluent dance coherent to the music.

858, TITLE: Faithful Extreme Rescaling via Generative Prior Reciprocated Invertible Representations
http://openaccess.thecvf.com//content/CVPR2022/html/Zhong_Faithful_Extreme_Rescaling_via_Generative_Prior_Reciprocated_Invertible_Representations_CVPR_2022_paper.html
AUTHORS: Zhixuan Zhong, Liangyu Chai, Yang Zhou, Bailin Deng, Jia Pan, Shengfeng He
HIGHLIGHT: This paper presents a Generative prior ReciprocAted Invertible rescaling Network (GRAIN) for generating faithful high-resolution (HR) images from low-resolution (LR) invertible images with an extreme upscaling factor (64x).

859, TITLE: Distillation Using Oracle Queries for Transformer-Based Human-Object Interaction Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Qu_Distillation_Using_Oracle_Queries_for_Transformer-Based_Human-Object_Interaction_Detection_CVPR_2022_paper.html
AUTHORS: Xian Qu, Changxing Ding, Xingao Li, Xubin Zhong, Dacheng Tao
HIGHLIGHT: To handle the first problem, we propose an efficient knowledge distillation model, named Distillation using Oracle Queries (DOQ), which shares parameters between teacher and student networks.

860, TITLE: Proto2Proto: Can You Recognize the Car, the Way I Do?
http://openaccess.thecvf.com//content/CVPR2022/html/Keswani_Proto2Proto_Can_You_Recognize_the_Car_the_Way_I_Do_CVPR_2022_paper.html
AUTHORS: Monish Keswani, Sriranjani Ramakrishnan, Nishant Reddy, Vineeth N Balasubramanian
HIGHLIGHT: We present Proto2Proto, a novel method to transfer interpretability of one prototypical part network to another via knowledge distillation.

861, TITLE: Learning Local-Global Contextual Adaptation for Multi-Person Pose Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Xue_Learning_Local-Global_Contextual_Adaptation_for_Multi-Person_Pose_Estimation_CVPR_2022_paper.html
AUTHORS: Nan Xue, Tianfu Wu, Gui-Song Xia, Liangpei Zhang
HIGHLIGHT: With a new and strong observation that the localization issue of the center-offset formulation can be remedied in a local-window search scheme in an ideal situation, we propose a multi-person pose estimation approach, dubbed as LOGO-CAP, by learning the LOcal-GlObal Contextual Adaptation for human Pose.

862, TITLE: Learning Video Representations of Human Motion From Synthetic Data
http://openaccess.thecvf.com//content/CVPR2022/html/Guo_Learning_Video_Representations_of_Human_Motion_From_Synthetic_Data_CVPR_2022_paper.html
AUTHORS: Xi Guo, Wei Wu, Dongliang Wang, Jing Su, Haisheng Su, Weihao Gan, Jian Huang, Qin Yang
HIGHLIGHT: In this paper, we take an early step towards video representation learning of human actions with the help of largescale synthetic videos, particularly for human motion representation enhancement.

863, TITLE: TVConv: Efficient Translation Variant Convolution for Layout-Aware Visual Processing
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_TVConv_Efficient_Translation_Variant_Convolution_for_Layout-Aware_Visual_Processing_CVPR_2022_paper.html
AUTHORS: Jierun Chen, Tianlang He, Weipeng Zhuo, Li Ma, Sangtae Ha, S.-H. Gary Chan
HIGHLIGHT: We observe that these applications naturally exhibit the characteristics of large intra-image (spatial) variance and small cross-image variance. This observation motivates our efficient translation variant convolution (TVConv) for layout-aware visual processing.

864, TITLE: Dual Adversarial Adaptation for Cross-Device Real-World Image Super-Resolution
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_Dual_Adversarial_Adaptation_for_Cross-Device_Real-World_Image_Super-Resolution_CVPR_2022_paper.html
AUTHORS: Xiaoqian Xu, Pengxu Wei, Weikai Chen, Yang Liu, Mingzhi Mao, Liang Lin, Guanbin Li
HIGHLIGHT: In this paper, we investigate a novel and practical task coded cross-device SR, which strives to adapt a real-world SR model trained on the paired images captured by one camera to low-resolution (LR) images captured by arbitrary target devices.

865, TITLE: FS6D: Few-Shot 6D Pose Estimation of Novel Objects
http://openaccess.thecvf.com//content/CVPR2022/html/He_FS6D_Few-Shot_6D_Pose_Estimation_of_Novel_Objects_CVPR_2022_paper.html
AUTHORS: Yisheng He, Yao Wang, Haoqiang Fan, Jian Sun, Qifeng Chen
HIGHLIGHT: In this work, we study a new open set problem; the few-shot 6D object poses estimation: estimating the 6D pose of an unknown object by a few support views without extra training.

866, TITLE: Habitat-Web: Learning Embodied Object-Search Strategies From Human Demonstrations at Scale
http://openaccess.thecvf.com//content/CVPR2022/html/Ramrakhya_Habitat-Web_Learning_Embodied_Object-Search_Strategies_From_Human_Demonstrations_at_Scale_CVPR_2022_paper.html
AUTHORS: Ram Ramrakhya, Eric Undersander, Dhruv Batra, Abhishek Das

HIGHLIGHT: We present a large-scale study of imitating human demonstrations on tasks that require a virtual robot to search for objects in new environments - (1) ObjectGoal Navigation (e.g. 'find & go to a chair') and (2) Pick&Place (e.g. 'find mug, pick mug, find counter, place mug on counter').

867, TITLE: The Probabilistic Normal Epipolar Constraint for Frame-to-Frame Rotation Optimization Under Uncertain Feature Positions
http://openaccess.thecvf.com//content/CVPR2022/html/Muhle_The_Probabilistic_Normal_Epipolar_Constraint_for_Frame-to-Frame_Rotation_Optimization_Under_CVPR_2022_paper.html
AUTHORS: Dominik Muhle, Lukas Koestler, Nikolaus Demmel, Florian Bernard, Daniel Cremers
HIGHLIGHT: However, their approach does not take into account uncertainties, so that the accuracy of the estimated relative pose is highly dependent on accurate feature positions in the target frame. In this work, we introduce the probabilistic normal epipolar constraint (PNEC) that overcomes this limitation by accounting for anisotropic and inhomogeneous uncertainties in the feature positions.

868, TITLE: Vision-Language Pre-Training for Boosting Scene Text Detectors
http://openaccess.thecvf.com//content/CVPR2022/html/Song_Vision-Language_Pre-Training_for_Boosting_Scene_Text_Detectors_CVPR_2022_paper.html
AUTHORS: Sibo Song, Jianqiang Wan, Zhibo Yang, Jun Tang, Wenqing Cheng, Xiang Bai, Cong Yao
HIGHLIGHT: In this paper, we specifically adapt vision-language joint learning for scene text detection, a task that intrinsically involves cross-modal interaction between the two modalities: vision and language, since text is the written form of language.

869, TITLE: Reflection and Rotation Symmetry Detection via Equivariant Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Seo_Reflection_and_Rotation_Symmetry_Detection_via_Equivariant_Learning_CVPR_2022_paper.html
AUTHORS: Ahyun Seo, Byungjin Kim, Suha Kwak, Minsu Cho
HIGHLIGHT: In this work, we introduce a group-equivariant convolutional network for symmetry detection, dubbed EquiSym, which leverages equivariant feature maps with respect to a dihedral group of reflection and rotation.

870, TITLE: BoostMIS: Boosting Medical Image Semi-Supervised Learning With Adaptive Pseudo Labeling and Informative Active Annotation
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_BoostMIS_Boosting_Medical_Image_Semi-Supervised_Learning_With_Adaptive_Pseudo_Labeling_CVPR_2022_paper.html
AUTHORS: Wenqiao Zhang, Lei Zhu, James Hallinan, Shengyu Zhang, Andrew Makmur, Qingpeng Cai, Beng Chin Ooi
HIGHLIGHT: In this paper, we propose a novel semi-supervised learning (SSL) framework named BoostMIS that combines adaptive pseudo labeling and informative active annotation to unleash the potential of medical image SSL models: (1) BoostMIS can adaptively leverage the cluster assumption and consistency regularization of the unlabeled data according to the current learning status.

871, TITLE: Simple but Effective: CLIP Embeddings for Embodied AI
http://openaccess.thecvf.com//content/CVPR2022/html/Khandelwal_Simple_but_Effective_CLIP_Embeddings_for_Embodied_AI_CVPR_2022_paper.html
AUTHORS: Apoorv Khandelwal, Luca Weihs, Roozbeh Mottaghi, Aniruddha Kembhavi
HIGHLIGHT: We investigate the effectiveness of CLIP visual backbones for Embodied AI tasks.

872, TITLE: NomMer: Nominate Synergistic Context in Vision Transformer for Visual Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_NomMer_Nominate_Synergistic_Context_in_Vision_Transformer_for_Visual_Recognition_CVPR_2022_paper.html
AUTHORS: Hao Liu, Xinghua Jiang, Xin Li, Zhimin Bao, Deqiang Jiang, Bo Ren
HIGHLIGHT: Alternatively, a more graceful way is that global and local context can adaptively contribute per se to accommodate different visual data. To achieve this goal, we in this paper propose a novel ViT architecture, termed NomMer, which can dynamically Nominate the synergistic global-local context in vision transforMer.

873, TITLE: HOI4D: A 4D Egocentric Dataset for Category-Level Human-Object Interaction
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_HOI4D_A_4D_Egocentric_Dataset_for_Category-Level_Human-Object_Interaction_CVPR_2022_paper.html
AUTHORS: Yunze Liu, Yun Liu, Che Jiang, Kangbo Lyu, Weikang Wan, Hao Shen, Boqiang Liang, Zhoujie Fu, He Wang, Li Yi
HIGHLIGHT: We present HOI4D, a large-scale 4D egocentric dataset with rich annotations, to catalyze the research of category-level human-object interaction.

874, TITLE: Collaborative Transformers for Grounded Situation Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Cho_Collaborative_Transformers_for_Grounded_Situation_Recognition_CVPR_2022_paper.html
AUTHORS: Junhyeong Cho, Youngseok Yoon, Suha Kwak
HIGHLIGHT: Grounded situation recognition is the task of predicting the main activity, entities playing certain roles within the activity, and bounding-box groundings of the entities in the given image. To effectively deal with this challenging task, we introduce a novel approach where the two processes for activity classification and entity estimation are interactive and complementary.

875, TITLE: DyRep: Bootstrapping Training With Dynamic Re-Parameterization
http://openaccess.thecvf.com//content/CVPR2022/html/Huang_DyRep_Bootstrapping_Training_With_Dynamic_Re-Parameterization_CVPR_2022_paper.html
AUTHORS: Tao Huang, Shan You, Bohan Zhang, Yuxuan Du, Fei Wang, Chen Qian, Chang Xu
HIGHLIGHT: As such, the price to pay is an expensive computational overhead to manipulate these unnecessary behaviors. To eliminate the above caveats, we aim to bootstrap the training with minimal cost by devising a dynamic re-parameterization (DyRep) method, which encodes Rep technique into the training process that dynamically evolves the network structures.

876, TITLE: Not All Labels Are Equal: Rationalizing the Labeling Costs for Training Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Elezi_Not_All_Labels_Are_Equal_Rationalizing_the_Labeling_Costs_for_CVPR_2022_paper.html
AUTHORS: Ismail Elezi, Zhiding Yu, Anima Anandkumar, Laura Leal-Taixé, Jose M. Alvarez
HIGHLIGHT: In this work, we propose a unified framework for active learning, that considers both the uncertainty and the robustness of the detector, ensuring that the network performs well in all classes.

877, TITLE: CPPF: Towards Robust Category-Level 9D Pose Estimation in the Wild
http://openaccess.thecvf.com//content/CVPR2022/html/You_CPPF_Towards_Robust_Category-Level_9D_Pose_Estimation_in_the_Wild_CVPR_2022_paper.html
AUTHORS: Yang You, Ruoxi Shi, Weiming Wang, Cewu Lu
HIGHLIGHT: In this paper, we tackle the problem of category-level 9D pose estimation in the wild, given a single RGB-D frame.

878, TITLE: Interact Before Align: Leveraging Cross-Modal Knowledge for Domain Adaptive Action Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Interact_Before_Align_Leveraging_Cross-Modal_Knowledge_for_Domain_Adaptive_Action_CVPR_2022_paper.html
AUTHORS: Lijin Yang, Yifei Huang, Yusuke Sugano, Yoichi Sato
HIGHLIGHT: Cross-modal knowledge interaction allows other modalities to supplement missing transferable information because of the cross-modal complementarity. Also, the most transferable aspects of data can be highlighted using cross-modal consensus. In this work, we present a novel model that jointly considers these two characteristics for domain adaptive action recognition.

879, TITLE: Interactive Disentanglement: Learning Concepts by Interacting With Their Prototype Representations
http://openaccess.thecvf.com//content/CVPR2022/html/Stammer_Interactive_Disentanglement_Learning_Concepts_by_Interacting_With_Their_Prototype_Representations_CVPR_2022_paper.html
AUTHORS: Wolfgang Stammer, Marius Memmel, Patrick Schramowski, Kristian Kersting
HIGHLIGHT: In this work, we show the advantages of prototype representations for understanding and revising the latent space of neural concept learners.

880, TITLE: CDGNet: Class Distribution Guided Network for Human Parsing
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_CDGNet_Class_Distribution_Guided_Network_for_Human_Parsing_CVPR_2022_paper.html
AUTHORS: Kunliang Liu, Ouk Choi, Jianming Wang, Wonjun Hwang
HIGHLIGHT: Probably, a human head is less likely to be under the feet, and arms are more likely to be near the torso. Inspired by this observation, we make instance class distributions by accumulating the original human parsing label in the horizontal and vertical directions, which can be utilized as supervision signals.

881, TITLE: Recall@k Surrogate Loss With Large Batches and Similarity Mixup
http://openaccess.thecvf.com//content/CVPR2022/html/Patel_Recallk_Surrogate_Loss_With_Large_Batches_and_Similarity_Mixup_CVPR_2022_paper.html
AUTHORS: Yash Patel, Giorgos Tolias, JiÅí Matas
HIGHLIGHT: A differentiable surrogate loss for the recall is proposed in this work.

882, TITLE: Direct Voxel Grid Optimization: Super-Fast Convergence for Radiance Fields Reconstruction
http://openaccess.thecvf.com//content/CVPR2022/html/Sun_Direct_Voxel_Grid_Optimization_Super-Fast_Convergence_for_Radiance_Fields_Reconstruction_CVPR_2022_paper.html
AUTHORS: Cheng Sun, Min Sun, Hwann-Tzong Chen
HIGHLIGHT: We present a super-fast convergence approach to reconstructing the per-scene radiance field from a set of images that capture the scene with known poses.

883, TITLE: Continual Test-Time Domain Adaptation
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Continual_Test-Time_Domain_Adaptation_CVPR_2022_paper.html
AUTHORS: Qin Wang, Olga Fink, Luc Van Gool, Dengxin Dai
HIGHLIGHT: The noisy pseudo-labels can further lead to error accumulation and catastrophic forgetting. To tackle these issues, we propose a continual test-time adaptation approach (CoTTA) which comprises two parts.

884, TITLE: URetinex-Net: Retinex-Based Deep Unfolding Network for Low-Light Image Enhancement
http://openaccess.thecvf.com//content/CVPR2022/html/Wu_URetinex-Net_Retinex-Based_Deep_Unfolding_Network_for_Low-Light_Image_Enhancement_CVPR_2022_paper.html
AUTHORS: Wenhui Wu, Jian Weng, Pingping Zhang, Xu Wang, Wenhan Yang, Jianmin Jiang

94

HIGHLIGHT: However, the commonly used hand-crafted priors and optimization-driven solutions lead to the absence of adaptivity and efficiency. To address these issues, in this paper, we propose a Retinex-based deep unfolding network (URetinex-Net), which unfolds an optimization problem into a learnable network to decompose a low-light image into reflectance and illumination layers.

885, TITLE: Towards Multi-Domain Single Image Dehazing via Test-Time Training
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Towards_Multi-Domain_Single_Image_Dehazing_via_Test-Time_Training_CVPR_2022_paper.html
AUTHORS: Huan Liu, Zijun Wu, Liangyan Li, Sadaf Salehkalaibar, Jun Chen, Keyan Wang
HIGHLIGHT: However, we observe that this training strategy tends to compromise the model performance on individual datasets. Motivated by this observation, we propose a test-time training method which leverages a helper network to assist the dehazing model in better adapting to a domain of interest.

886, TITLE: Vox2Cortex: Fast Explicit Reconstruction of Cortical Surfaces From 3D MRI Scans With Geometric Deep Neural Networks
http://openaccess.thecvf.com//content/CVPR2022/html/Bongratz_Vox2Cortex_Fast_Explicit_Reconstruction_of_Cortical_Surfaces_From_3D_MRI_CVPR_2022_paper.html
AUTHORS: Fabian Bongratz, Anne-Marie Rickmann, Sebastian Pösterl, Christian Wachinger
HIGHLIGHT: Although traditional and deep learning-based algorithmic pipelines exist for this purpose, they have two major drawbacks: lengthy runtimes of multiple hours (traditional) or intricate post-processing, such as mesh extraction and topology correction (deep learning-based). In this work, we address both of these issues and propose Vox2Cortex, a deep learning-based algorithm that directly yields topologically correct, three-dimensional meshes of the boundaries of the cortex.

887, TITLE: Deep Safe Multi-View Clustering: Reducing the Risk of Clustering Performance Degradation Caused by View Increase
http://openaccess.thecvf.com//content/CVPR2022/html/Tang_Deep_Safe_Multi-View_Clustering_Reducing_the_Risk_of_Clustering_Performance_CVPR_2022_paper.html
AUTHORS: Huayi Tang, Yong Liu
HIGHLIGHT: However, we observe that learning from data with more views is not guaranteed to achieve better clustering performance than from data with fewer views. To address this issue, we propose a general deep learning based framework that is guaranteed to reduce the risk of performance degradation caused by view increase.

888, TITLE: Dynamic MLP for Fine-Grained Image Classification by Leveraging Geographical and Temporal Information
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Dynamic_MLP_for_Fine-Grained_Image_Classification_by_Leveraging_Geographical_and_CVPR_2022_paper.html
AUTHORS: Lingfeng Yang, Xiang Li, Renjie Song, Borui Zhao, Juntian Tao, Shihao Zhou, Jiajun Liang, Jian Yang
HIGHLIGHT: To fully explore the potential of multimodal information, we propose a dynamic MLP on top of the image representation, which interacts with multimodal features at a higher and broader dimension.

889, TITLE: HP-Capsule: Unsupervised Face Part Discovery by Hierarchical Parsing Capsule Network
http://openaccess.thecvf.com//content/CVPR2022/html/Yu_HP-Capsule_Unsupervised_Face_Part_Discovery_by_Hierarchical_Parsing_Capsule_Network_CVPR_2022_paper.html
AUTHORS: Chang Yu, Xiangyu Zhu, Xiaomei Zhang, Zidu Wang, Zhaoxiang Zhang, Zhen Lei
HIGHLIGHT: In this paper, we propose a Hierarchical Parsing Capsule Network (HP-Capsule) for unsupervised face subpart-part discovery.

890, TITLE: ScanQA: 3D Question Answering for Spatial Scene Understanding
http://openaccess.thecvf.com//content/CVPR2022/html/Azuma_ScanQA_3D_Question_Answering_for_Spatial_Scene_Understanding_CVPR_2022_paper.html
AUTHORS: Daichi Azuma, Taiki Miyanishi, Shuhei Kurita, Motoaki Kawanabe
HIGHLIGHT: We propose a baseline model for 3D-QA, named ScanQA model, where the model learns a fused descriptor from 3D object proposals and encoded sentence embeddings.

891, TITLE: MuKEA: Multimodal Knowledge Extraction and Accumulation for Knowledge-Based Visual Question Answering
http://openaccess.thecvf.com//content/CVPR2022/html/Ding_MuKEA_Multimodal_Knowledge_Extraction_and_Accumulation_for_Knowledge-Based_Visual_Question_CVPR_2022_paper.html
AUTHORS: Yang Ding, Jing Yu, Bang Liu, Yue Hu, Mingxin Cui, Qi Wu
HIGHLIGHT: In this paper, we propose MuKEA to represent multimodal knowledge by an explicit triplet to correlate visual objects and fact answers with implicit relations.

892, TITLE: Class-Incremental Learning by Knowledge Distillation With Adaptive Feature Consolidation
http://openaccess.thecvf.com//content/CVPR2022/html/Kang_Class-Incremental_Learning_by_Knowledge_Distillation_With_Adaptive_Feature_Consolidation_CVPR_2022_paper.html
AUTHORS: Minsoo Kang, Jaeyoo Park, Bohyung Han
HIGHLIGHT: We present a novel class incremental learning approach based on deep neural networks, which continually learns new tasks with limited memory for storing examples in the previous tasks.

893, TITLE: Learning Program Representations for Food Images and Cooking Recipes

http://openaccess.thecvf.com//content/CVPR2022/html/Papadopoulos_Learning_Program_Representations_for_Food_Images_and_Cooking_Recipes_CVPR_2022_paper.html
AUTHORS:		Dim P. Papadopoulos, Enrique Mora, Nadiia Chepurko, Kuan Wei Huang, Ferda Ofli, Antonio Torralba
HIGHLIGHT:		In this paper, we are interested in modeling a how-to instructional procedure, such as a cooking recipe, with a meaningful and rich high-level representation.

894, TITLE:		Bending Graphs: Hierarchical Shape Matching Using Gated Optimal Transport
http://openaccess.thecvf.com//content/CVPR2022/html/Saleh_Bending_Graphs_Hierarchical_Shape_Matching_Using_Gated_Optimal_Transport_CVPR_2022_paper.html
AUTHORS:		Mahdi Saleh, Shun-Cheng Wu, Luca Cosmo, Nassir Navab, Benjamin Busam, Federico Tombari
HIGHLIGHT:		In this work, we investigate a hierarchical learning design, to which we incorporate local patch-level information and global shape-level structures.

895, TITLE:		Transform-Retrieve-Generate: Natural Language-Centric Outside-Knowledge Visual Question Answering
http://openaccess.thecvf.com//content/CVPR2022/html/Gao_Transform-Retrieve-Generate_Natural_Language-Centric_Outside-Knowledge_Visual_Question_Answering_CVPR_2022_paper.html
AUTHORS:		Feng Gao, Qing Ping, Govind Thattai, Aishwarya Reganti, Ying Nian Wu, Prem Natarajan
HIGHLIGHT:		In this paper, we call for an alternative paradigm for the OK-VQA task, which transforms the image into plain text, so that we can enable knowledge passage retrieval, and generative question-answering in the natural language space.

896, TITLE:		Federated Learning With Position-Aware Neurons
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Federated_Learning_With_Position-Aware_Neurons_CVPR_2022_paper.html
AUTHORS:		Xin-Chun Li, Yi-Chu Xu, Shaoming Song, Bingshuai Li, Yinchuan Li, Yunfeng Shao, De-Chuan Zhan
HIGHLIGHT:		Hence, we propose Position-Aware Neurons (PANs) as an alternative, fusing position-related values (i.e., position encodings) into neuron outputs.

897, TITLE:		Fair Contrastive Learning for Facial Attribute Classification
http://openaccess.thecvf.com//content/CVPR2022/html/Park_Fair_Contrastive_Learning_for_Facial_Attribute_Classification_CVPR_2022_paper.html
AUTHORS:		Sungho Park, Jewook Lee, Pilhyeon Lee, Sunhee Hwang, Dohyung Kim, Hyeran Byun
HIGHLIGHT:		In this paper, we for the first time analyze unfairness caused by supervised contrastive learning and propose a new Fair Supervised Contrastive Loss (FSCL) for fair visual representation learning.

898, TITLE:		MDAN: Multi-Level Dependent Attention Network for Visual Emotion Analysis
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_MDAN_Multi-Level_Dependent_Attention_Network_for_Visual_Emotion_Analysis_CVPR_2022_paper.html
AUTHORS:		Liwen Xu, Zhengtao Wang, Bin Wu, Simon Lui
HIGHLIGHT:		In this paper, we present the Multi-level Dependent Attention Network (MDAN) with two branches, to leverage the emotion hierarchy and the correlation between different affective levels and semantic levels.

899, TITLE:		Nested Hyperbolic Spaces for Dimensionality Reduction and Hyperbolic NN Design
http://openaccess.thecvf.com//content/CVPR2022/html/Fan_Nested_Hyperbolic_Spaces_for_Dimensionality_Reduction_and_Hyperbolic_NN_Design_CVPR_2022_paper.html
AUTHORS:		Xiran Fan, Chun-Hao Yang, Baba C. Vemuri
HIGHLIGHT:		In this paper, we present a novel fully hyperbolic neural network which uses the concept of projections (embeddings) followed by an intrinsic aggregation and a nonlinearity all within the hyperbolic space.

900, TITLE:		BNUDC: A Two-Branched Deep Neural Network for Restoring Images From Under-Display Cameras
http://openaccess.thecvf.com//content/CVPR2022/html/Koh_BNUDC_A_Two-Branched_Deep_Neural_Network_for_Restoring_Images_From_CVPR_2022_paper.html
AUTHORS:		Jaihyun Koh, Jangho Lee, Sungroh Yoon
HIGHLIGHT:		We introduce a deep neural network with two branches to reverse each type of degradation, which is more effective than performing both restorations in a single forward network.

901, TITLE:		RGB-Depth Fusion GAN for Indoor Depth Completion
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_RGB-Depth_Fusion_GAN_for_Indoor_Depth_Completion_CVPR_2022_paper.html
AUTHORS:		Haowen Wang, Mingyuan Wang, Zhengping Che, Zhiyuan Xu, Xiuquan Qiao, Mengshi Qi, Feifei Feng, Jian Tang
HIGHLIGHT:		In this paper, we design a novel two-branch end-to-end fusion network, which takes a pair of RGB and incompleted depth images as input to predict a dense and completed depth map.

902, TITLE:		Training Object Detectors From Scratch: An Empirical Study in the Era of Vision Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Hong_Training_Object_Detectors_From_Scratch_An_Empirical_Study_in_the_CVPR_2022_paper.html
AUTHORS:		Weixiang Hong, Jiangwei Lao, Wang Ren, Jian Wang, Jingdong Chen, Wei Chu
HIGHLIGHT:		Instead of proposing a specific vision transformer based detector, in this work, our goal is to reveal the insights of training vision transformer based detectors from scratch.

903, TITLE: RCL: Recurrent Continuous Localization for Temporal Action Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_RCL_Recurrent_Continuous_Localization_for_Temporal_Action_Detection_CVPR_2022_paper.html
AUTHORS: Qiang Wang, Yanhao Zhang, Yun Zheng, Pan Pan
HIGHLIGHT: State-of-the-art methods mostly rely on a dense anchoring scheme, where anchors are sampled uniformly over the temporal domain with a discretized grid, and then regress the accurate boundaries. In this paper, we revisit this foundational stage and introduce Recurrent Continuous Localization (RCL), which learns a fully continuous anchoring representation.

904, TITLE: C2SLR: Consistency-Enhanced Continuous Sign Language Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Zuo_C2SLR_Consistency-Enhanced_Continuous_Sign_Language_Recognition_CVPR_2022_paper.html
AUTHORS: Ronglai Zuo, Brian Mak
HIGHLIGHT: In this work, we propose two auxiliary constraints to enhance the CSLR backbones from the perspective of consistency.

905, TITLE: Human Trajectory Prediction With Momentary Observation
http://openaccess.thecvf.com//content/CVPR2022/html/Sun_Human_Trajectory_Prediction_With_Momentary_Observation_CVPR_2022_paper.html
AUTHORS: Jianhua Sun, Yuxuan Li, Liang Chai, Hao-Shu Fang, Yong-Lu Li, Cewu Lu
HIGHLIGHT: In this paper, we study a task named momentary trajectory prediction, which reduces the observed history from a long time sequence to an extreme situation of two frames, one frame for social and scene contexts and both frames for the velocity of agents.

906, TITLE: FoggyStereo: Stereo Matching With Fog Volume Representation
http://openaccess.thecvf.com//content/CVPR2022/html/Yao_FoggyStereo_Stereo_Matching_With_Fog_Volume_Representation_CVPR_2022_paper.html
AUTHORS: Chengtang Yao, Lidong Yu
HIGHLIGHT: Prior methods deem the fog as noise and discard it before matching. Different from them, we propose to explore depth hints from fog and improve stereo matching via these hints.

907, TITLE: Trajectory Optimization for Physics-Based Reconstruction of 3D Human Pose From Monocular Video
http://openaccess.thecvf.com//content/CVPR2022/html/Gartner_Trajectory_Optimization_for_Physics-Based_Reconstruction_of_3D_Human_Pose_From_CVPR_2022_paper.html
AUTHORS: Erik Gärtner, Mykhaylo Andriluka, Hongyi Xu, Cristian Sminchisescu
HIGHLIGHT: We focus on the task of estimating a physically plausible articulated human motion from monocular video.

908, TITLE: Directional Self-Supervised Learning for Heavy Image Augmentations
http://openaccess.thecvf.com//content/CVPR2022/html/Bai_Directional_Self-Supervised_Learning_for_Heavy_Image_Augmentations_CVPR_2022_paper.html
AUTHORS: Yalong Bai, Yifan Yang, Wei Zhang, Tao Mei
HIGHLIGHT: In this paper, we propose a directional self-supervised learning paradigm (DSSL), which is compatible with significantly more augmentations.

909, TITLE: Lifelong Unsupervised Domain Adaptive Person Re-Identification With Coordinated Anti-Forgetting and Adaptation
http://openaccess.thecvf.com//content/CVPR2022/html/Huang_Lifelong_Unsupervised_Domain_Adaptive_Person_Re-Identification_With_Coordinated_Anti-Forgetting_and_CVPR_2022_paper.html
AUTHORS: Zhipeng Huang, Zhizheng Zhang, Cuiling Lan, Wenjun Zeng, Peng Chu, Quanzeng You, Jiang Wang, Zicheng Liu, Zheng-Jun Zha
HIGHLIGHT: In this paper, to address more practical scenarios, we propose a new task, Lifelong Unsupervised Domain Adaptive (LUDA) person ReID.

910, TITLE: No-Reference Point Cloud Quality Assessment via Domain Adaptation
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_No-Reference_Point_Cloud_Quality_Assessment_via_Domain_Adaptation_CVPR_2022_paper.html
AUTHORS: Qi Yang, Yipeng Liu, Siheng Chen, Yiling Xu, Jun Sun
HIGHLIGHT: We present a novel no-reference quality assessment metric, the image transferred point cloud quality assessment (IT-PCQA), for 3D point clouds.

911, TITLE: Generating Representative Samples for Few-Shot Classification
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_Generating_Representative_Samples_for_Few-Shot_Classification_CVPR_2022_paper.html
AUTHORS: Jingyi Xu, Hieu Le
HIGHLIGHT: Few-shot class representations are often biased due to data scarcity. To mitigate this issue, we propose to generate visual samples based on semantic embeddings using a conditional variational autoencoder (CVAE) model.

912, TITLE: Comprehending and Ordering Semantics for Image Captioning
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Comprehending_and_Ordering_Semantics_for_Image_Captioning_CVPR_2022_paper.html
AUTHORS: Yehao Li, Yingwei Pan, Ting Yao, Tao Mei

HIGHLIGHT:	In this paper, we propose a new recipe of Transformer-style structure, namely Comprehending and Ordering Semantics Networks (COS-Net), that novelly unifies an enriched semantic comprehending and a learnable semantic ordering processes into a single architecture.

913, TITLE:	Dynamic Scene Graph Generation via Anticipatory Pre-Training
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Dynamic_Scene_Graph_Generation_via_Anticipatory_Pre-Training_CVPR_2022_paper.html
AUTHORS:	Yiming Li, Xiaoshan Yang, Changsheng Xu
HIGHLIGHT:	Inspired by the ability of humans to infer the visual relationship, we propose a novel anticipatory pre-training paradigm based on Transformer to explicitly model the temporal correlation of visual relationships in different frames to improve dynamic scene graph generation.

914, TITLE:	A Large-Scale Comprehensive Dataset and Copy-Overlap Aware Evaluation Protocol for Segment-Level Video Copy Detection
http://openaccess.thecvf.com//content/CVPR2022/html/He_A_Large-Scale_Comprehensive_Dataset_and_Copy-Overlap_Aware_Evaluation_Protocol_for_CVPR_2022_paper.html
AUTHORS:	Sifeng He, Xudong Yang, Chen Jiang, Gang Liang, Wei Zhang, Tan Pan, Qing Wang, Furong Xu, Chunguang Li, JinXiong Liu, Hui Xu, Kaiming Huang, Yuan Cheng, Feng Qian, Xiaobo Zhang, Lei Yang
HIGHLIGHT:	In this paper, we introduce VCSL (Video Copy Segment Localization), a new comprehensive segment-level annotated video copy dataset.

915, TITLE:	GaTector: A Unified Framework for Gaze Object Prediction
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_GaTector_A_Unified_Framework_for_Gaze_Object_Prediction_CVPR_2022_paper.html
AUTHORS:	Binglu Wang, Tao Hu, Baoshan Li, Xiaojuan Chen, Zhijie Zhang
HIGHLIGHT:	In this paper, we build a novel framework named GaTector to tackle the gaze object prediction problem in a unified way.

916, TITLE:	ELIC: Efficient Learned Image Compression With Unevenly Grouped Space-Channel Contextual Adaptive Coding
http://openaccess.thecvf.com//content/CVPR2022/html/He_ELIC_Efficient_Learned_Image_Compression_With_Unevenly_Grouped_Space-Channel_Contextual_CVPR_2022_paper.html
AUTHORS:	Dailan He, Ziming Yang, Weikun Peng, Rui Ma, Hongwei Qin, Yan Wang
HIGHLIGHT:	Combining the proposed uneven grouping model with existing context models, we obtain a spatial-channel contextual adaptive model to improve the coding performance without damage to running speed. Then we study the structure of the main transform and propose an efficient model, ELIC, to achieve state-of-the-art speed and compression ability.

917, TITLE:	CSWin Transformer: A General Vision Transformer Backbone With Cross-Shaped Windows
http://openaccess.thecvf.com//content/CVPR2022/html/Dong_CSWin_Transformer_A_General_Vision_Transformer_Backbone_With_Cross-Shaped_Windows_CVPR_2022_paper.html
AUTHORS:	Xiaoyi Dong, Jianmin Bao, Dongdong Chen, Weiming Zhang, Nenghai Yu, Lu Yuan, Dong Chen, Baining Guo
HIGHLIGHT:	We present CSWin Transformer, an efficient and effective Transformer-based backbone for general-purpose vision tasks.

918, TITLE:	LaTr: Layout-Aware Transformer for Scene-Text VQA
http://openaccess.thecvf.com//content/CVPR2022/html/Biten_LaTr_Layout-Aware_Transformer_for_Scene-Text_VQA_CVPR_2022_paper.html
AUTHORS:	Ali Furkan Biten, Ron Litman, Yusheng Xie, Srikar Appalaraju, R. Manmatha
HIGHLIGHT:	We propose a novel multimodal architecture for Scene Text Visual Question Answering (STVQA), named Layout-Aware Transformer (LaTr).

919, TITLE:	Label Relation Graphs Enhanced Hierarchical Residual Network for Hierarchical Multi-Granularity Classification
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Label_Relation_Graphs_Enhanced_Hierarchical_Residual_Network_for_Hierarchical_Multi-Granularity_CVPR_2022_paper.html
AUTHORS:	Jingzhou Chen, Peng Wang, Jian Liu, Yuntao Qian
HIGHLIGHT:	In this paper, we study the HMC problem in which objects are labeled at any level of the hierarchy.

920, TITLE:	ITSA: An Information-Theoretic Approach to Automatic Shortcut Avoidance and Domain Generalization in Stereo Matching Networks
http://openaccess.thecvf.com//content/CVPR2022/html/Chuah_ITSA_An_Information-Theoretic_Approach_to_Automatic_Shortcut_Avoidance_and_Domain_CVPR_2022_paper.html
AUTHORS:	WeiQin Chuah, Ruwan Tennakoon, Reza Hoseinnezhad, Alireza Bab-Hadiashar, David Suter
HIGHLIGHT:	In this paper, we attempt to unfold an important factor that hinders the networks from generalizing across domains: through the lens of shortcut learning.

921, TITLE:	Enhancing Face Recognition With Self-Supervised 3D Reconstruction
http://openaccess.thecvf.com//content/CVPR2022/html/He_Enhancing_Face_Recognition_With_Self-Supervised_3D_Reconstruction_CVPR_2022_paper.html
AUTHORS:	Mingjie He, Jie Zhang, Shiguang Shan, Xilin Chen

HIGHLIGHT: In this paper, we propose to enhance face recognition with a bypass of self-supervised 3D reconstruction, which enforces the neural backbone to focus on the identity-related depth and albedo information while neglects the identity-irrelevant pose and illumination information.

922, TITLE: HeadNeRF: A Real-Time NeRF-Based Parametric Head Model
http://openaccess.thecvf.com//content/CVPR2022/html/Hong_HeadNeRF_A_Real-Time_NeRF-Based_Parametric_Head_Model_CVPR_2022_paper.html
AUTHORS: Yang Hong, Bo Peng, Haiyao Xiao, Ligang Liu, Juyong Zhang
HIGHLIGHT: In this paper, we propose HeadNeRF, a novel NeRF-based parametric head model that integrates the neural radiance field to the parametric representation of the human head.

923, TITLE: FvOR: Robust Joint Shape and Pose Optimization for Few-View Object Reconstruction
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_FvOR_Robust_Joint_Shape_and_Pose_Optimization_for_Few-View_Object_CVPR_2022_paper.html
AUTHORS: Zhenpei Yang, Zhile Ren, Miguel Angel Bautista, Zaiwei Zhang, Qi Shan, Qixing Huang
HIGHLIGHT: In this paper, we present FvOR, a learning-based object reconstruction method that predicts accurate 3D models given a few images with noisy input poses.

924, TITLE: Reduce Information Loss in Transformers for Pluralistic Image Inpainting
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Reduce_Information_Loss_in_Transformers_for_Pluralistic_Image_Inpainting_CVPR_2022_paper.html
AUTHORS: Qiankun Liu, Zhentao Tan, Dongdong Chen, Qi Chu, Xiyang Dai, Yinpeng Chen, Mengchen Liu, Lu Yuan, Nenghai Yu
HIGHLIGHT: To keep input information as much as possible, we propose a new transformer based framework "PUT".

925, TITLE: Replacing Labeled Real-Image Datasets With Auto-Generated Contours
http://openaccess.thecvf.com//content/CVPR2022/html/Kataoka_Replacing_Labeled_Real-Image_Datasets_With_Auto-Generated_Contours_CVPR_2022_paper.html
AUTHORS: Hirokatsu Kataoka, Ryo Hayamizu, Ryosuke Yamada, Kodai Nakashima, Sora Takashima, Xinyu Zhang, Edgar Josafat Martinez-Noriega, Nakamasa Inoue, Rio Yokota
HIGHLIGHT: In the present work, we show that the performance of formula-driven supervised learning (FDSL) can match or even exceed that of ImageNet-21k without the use of real images, human-, and self-supervision during the pre-training of Vision Transformers (ViTs).

926, TITLE: Cross-Modal Transferable Adversarial Attacks From Images to Videos
http://openaccess.thecvf.com//content/CVPR2022/html/Wei_Cross-Modal_Transferable_Adversarial_Attacks_From_Images_to_Videos_CVPR_2022_paper.html
AUTHORS: Zhipeng Wei, Jingjing Chen, Zuxuan Wu, Yu-Gang Jiang
HIGHLIGHT: Specifically, motivated by the observation that the low-level feature space between images and video frames are similar, we propose a simple yet effective cross-modal attack method, named as Image To Video (I2V) attack.

927, TITLE: Few Could Be Better Than All: Feature Sampling and Grouping for Scene Text Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Tang_Few_Could_Be_Better_Than_All_Feature_Sampling_and_Grouping_CVPR_2022_paper.html
AUTHORS: Jingqun Tang, Wenqing Zhang, Hongye Liu, MingKun Yang, Bo Jiang, Guanglong Hu, Xiang Bai
HIGHLIGHT: In this paper, we present a simple yet effective transformer-based architecture for scene text detection.

928, TITLE: Do Explanations Explain? Model Knows Best
http://openaccess.thecvf.com//content/CVPR2022/html/Khakzar_Do_Explanations_Explain_Model_Knows_Best_CVPR_2022_paper.html
AUTHORS: Ashkan Khakzar, Pedram Khorsandi, Rozhin Nobahari, Nassir Navab
HIGHLIGHT: Thus we propose an empirical framework for axiomatic evaluation of explanation methods.

929, TITLE: WebQA: Multihop and Multimodal QA
http://openaccess.thecvf.com//content/CVPR2022/html/Chang_WebQA_Multihop_and_Multimodal_QA_CVPR_2022_paper.html
AUTHORS: Yingshan Chang, Mridu Narang, Hisami Suzuki, Guihong Cao, Jianfeng Gao, Yonatan Bisk
HIGHLIGHT: In this work, we introduce WebQA, a challenging new benchmark that proves difficult for large-scale state-of-the-art models which lack language groundable visual representations for novel objects and the ability to reason, yet trivial for humans.

930, TITLE: Occlusion-Robust Face Alignment Using a Viewpoint-Invariant Hierarchical Network Architecture
http://openaccess.thecvf.com//content/CVPR2022/html/Zhu_Occlusion-Robust_Face_Alignment_Using_a_Viewpoint-Invariant_Hierarchical_Network_Architecture_CVPR_2022_paper.html
AUTHORS: Congcong Zhu, Xintong Wan, Shaorong Xie, Xiaoqiang Li, Yinzheng Gu
HIGHLIGHT: This paper proposes a new network architecture called GlomFace to model the facial hierarchies against various occlusions, which draws inspiration from the viewpoint-invariant hierarchy of facial structure.

931, TITLE: BasicVSR++: Improving Video Super-Resolution With Enhanced Propagation and Alignment

http://openaccess.thecvf.com//content/CVPR2022/html/Chan_BasicVSR_Improving_Video_Super-Resolution_With_Enhanced_Propagation_and_Alignment_CVPR_2022_paper.html
AUTHORS:　　　Kelvin C.K. Chan, Shangchen Zhou, Xiangyu Xu, Chen Change Loy
HIGHLIGHT:　　　In this study, we redesign BasicVSR by proposing second-order grid propagation and flow-guided deformable alignment.

932, TITLE:　　　IDR: Self-Supervised Image Denoising via Iterative Data Refinement
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_IDR_Self-Supervised_Image_Denoising_via_Iterative_Data_Refinement_CVPR_2022_paper.html
AUTHORS:　　　Yi Zhang, Dasong Li, Ka Lung Law, Xiaogang Wang, Hongwei Qin, Hongsheng Li
HIGHLIGHT:　　　In this paper, we present a practical unsupervised image denoising method to achieve state-of-the-art denoising performance.

933, TITLE:　　　MogFace: Towards a Deeper Appreciation on Face Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_MogFace_Towards_a_Deeper_Appreciation_on_Face_Detection_CVPR_2022_paper.html
AUTHORS:　　　Yang Liu, Fei Wang, Jiankang Deng, Zhipeng Zhou, Baigui Sun, Hao Li
HIGHLIGHT:　　　In this paper, we focus on resolving three aforementioned challenges that exiting methods are difficult to finish off and present a novel face detector, termed MogFace.

934, TITLE:　　　GuideFormer: Transformers for Image Guided Depth Completion
http://openaccess.thecvf.com//content/CVPR2022/html/Rho_GuideFormer_Transformers_for_Image_Guided_Depth_Completion_CVPR_2022_paper.html
AUTHORS:　　　Kyeongha Rho, Jinsung Ha, Youngjung Kim
HIGHLIGHT:　　　In this paper, we propose GuideFormer, a fully transformer-based architecture for dense depth completion.

935, TITLE:　　　Multi-Label Iterated Learning for Image Classification With Label Ambiguity
http://openaccess.thecvf.com//content/CVPR2022/html/Rajeswar_Multi-Label_Iterated_Learning_for_Image_Classification_With_Label_Ambiguity_CVPR_2022_paper.html
AUTHORS:　　　Sai Rajeswar, Pau Rodríguez, Soumye Singhal, David Vazquez, Aaron Courville
HIGHLIGHT:　　　Inspired by language emergence literature, we propose multi-label iterated learning (MILe) to incorporate the inductive biases of multi-label learning from single labels using the framework of iterated learning.

936, TITLE:　　　Region-Aware Face Swapping
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_Region-Aware_Face_Swapping_CVPR_2022_paper.html
AUTHORS:　　　Chao Xu, Jiangning Zhang, Miao Hua, Qian He, Zili Yi, Yong Liu
HIGHLIGHT:　　　This paper presents a novel Region-Aware Face Swapping (RAFSwap) network to achieve identity-consistent harmonious high-resolution face generation in a local-global manner: 1) Local Facial Region-Aware (FRA) branch augments local identity-relevant features by introducing the Transformer to effectively model misaligned cross-scale semantic interaction.

937, TITLE:　　　Towards Language-Free Training for Text-to-Image Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Zhou_Towards_Language-Free_Training_for_Text-to-Image_Generation_CVPR_2022_paper.html
AUTHORS:　　　Yufan Zhou, Ruiyi Zhang, Changyou Chen, Chunyuan Li, Chris Tensmeyer, Tong Yu, Jiuxiang Gu, Jinhui Xu, Tong Sun
HIGHLIGHT:　　　In this paper, we propose the first work to train text-to-image generation models without any text data.

938, TITLE:　　　Learning Affinity From Attention: End-to-End Weakly-Supervised Semantic Segmentation With Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Ru_Learning_Affinity_From_Attention_End-to-End_Weakly-Supervised_Semantic_Segmentation_With_Transformers_CVPR_2022_paper.html
AUTHORS:　　　Lixiang Ru, Yibing Zhan, Baosheng Yu, Bo Du
HIGHLIGHT:　　　In this paper, to address the aforementioned problem, we introduce Transformers, which naturally integrate global information, to generate more integral initial pseudo labels for end-to-end WSSS.

939, TITLE:　　　Pushing the Envelope of Gradient Boosting Forests via Globally-Optimized Oblique Trees
http://openaccess.thecvf.com//content/CVPR2022/html/Gabidolla_Pushing_the_Envelope_of_Gradient_Boosting_Forests_via_Globally-Optimized_Oblique_CVPR_2022_paper.html
AUTHORS:　　　Magzhan Gabidolla, Miguel Á. Carreira-Perpiñán
HIGHLIGHT:　　　To date, all successful GB versions use axis-aligned trees trained in a suboptimal way via greedy recursive partitioning. We address this gap by using a more powerful type of trees (having hyperplane splits) and an algorithm that can optimize, globally over all the tree parameters, the objective function that GB dictates.

940, TITLE:　　　Physical Simulation Layer for Accurate 3D Modeling
http://openaccess.thecvf.com//content/CVPR2022/html/Mezghanni_Physical_Simulation_Layer_for_Accurate_3D_Modeling_CVPR_2022_paper.html
AUTHORS:　　　Mariem Mezghanni, Théo Bodrito, Malika Boulkenafed, Maks Ovsjanikov
HIGHLIGHT:　　　We introduce a novel approach for generative 3D modeling that explicitly encourages the physical and thus functional consistency of the generated shapes.

941, TITLE:　　　Deformable Sprites for Unsupervised Video Decomposition

http://openaccess.thecvf.com//content/CVPR2022/html/Ye_Deformable_Sprites_for_Unsupervised_Video_Decomposition_CVPR_2022_paper.html

AUTHORS: Vickie Ye, Zhengqi Li, Richard Tucker, Angjoo Kanazawa, Noah Snavely

HIGHLIGHT: We describe a method to extract persistent elements of a dynamic scene from an input video.

942, TITLE: CamLiFlow: Bidirectional Camera-LiDAR Fusion for Joint Optical Flow and Scene Flow Estimation

http://openaccess.thecvf.com//content/CVPR2022/html/Liu_CamLiFlow_Bidirectional_Camera-LiDAR_Fusion_for_Joint_Optical_Flow_and_Scene_CVPR_2022_paper.html

AUTHORS: Haisong Liu, Tao Lu, Yihui Xu, Jia Liu, Wenjie Li, Lijun Chen

HIGHLIGHT: In this paper, we study the problem of jointly estimating the optical flow and scene flow from synchronized 2D and 3D data.

943, TITLE: FERV39k: A Large-Scale Multi-Scene Dataset for Facial Expression Recognition in Videos

http://openaccess.thecvf.com//content/CVPR2022/html/Wang_FERV39k_A_Large-Scale_Multi-Scene_Dataset_for_Facial_Expression_Recognition_in_CVPR_2022_paper.html

AUTHORS: Yan Wang, Yixuan Sun, Yiwen Huang, Zhongying Liu, Shuyong Gao, Wei Zhang, Weifeng Ge, Wenqiang Zhang

HIGHLIGHT: For example, the "Happy" expression with high intensity in Talk-Show is more discriminating than the same expression with low intensity in Official-Event. To fill this gap, we build a large-scale multi-scene dataset, coined as FERV39k.

944, TITLE: Learning To Detect Mobile Objects From LiDAR Scans Without Labels

http://openaccess.thecvf.com//content/CVPR2022/html/You_Learning_To_Detect_Mobile_Objects_From_LiDAR_Scans_Without_Labels_CVPR_2022_paper.html

AUTHORS: Yurong You, Katie Luo, Cheng Perng Phoo, Wei-Lun Chao, Wen Sun, Bharath Hariharan, Mark Campbell, Kilian Q. Weinberger

HIGHLIGHT: This paper proposes an alternative approach entirely based on unlabeled data, which can be collected cheaply and in abundance almost everywhere on earth.

945, TITLE: BNV-Fusion: Dense 3D Reconstruction Using Bi-Level Neural Volume Fusion

http://openaccess.thecvf.com//content/CVPR2022/html/Li_BNV-Fusion_Dense_3D_Reconstruction_Using_Bi-Level_Neural_Volume_Fusion_CVPR_2022_paper.html

AUTHORS: Kejie Li, Yansong Tang, Victor Adrian Prisacariu, Philip H.S. Torr

HIGHLIGHT: In order to incrementally integrate new depth maps into a global neural implicit representation, we propose a novel bi-level fusion strategy that considers both efficiency and reconstruction quality by design.

946, TITLE: Probabilistic Representations for Video Contrastive Learning

http://openaccess.thecvf.com//content/CVPR2022/html/Park_Probabilistic_Representations_for_Video_Contrastive_Learning_CVPR_2022_paper.html

AUTHORS: Jungin Park, Jiyoung Lee, Ig-Jae Kim, Kwanghoon Sohn

HIGHLIGHT: This paper presents Probabilistic Video Contrastive Learning, a self-supervised representation learning method that bridges contrastive learning with probabilistic representation.

947, TITLE: EnvEdit: Environment Editing for Vision-and-Language Navigation

http://openaccess.thecvf.com//content/CVPR2022/html/Li_EnvEdit_Environment_Editing_for_Vision-and-Language_Navigation_CVPR_2022_paper.html

AUTHORS: Jialu Li, Hao Tan, Mohit Bansal

HIGHLIGHT: Due to limited available data for agent training and finite diversity in navigation environments, it is challenging for the agent to generalize to new, unseen environments. To address this problem, we propose EnvEdit, a data augmentation method that creates new environments by editing existing environments, which are used to train a more generalizable agent.

948, TITLE: Omnivore: A Single Model for Many Visual Modalities

http://openaccess.thecvf.com//content/CVPR2022/html/Girdhar_Omnivore_A_Single_Model_for_Many_Visual_Modalities_CVPR_2022_paper.html

AUTHORS: Rohit Girdhar, Mannat Singh, Nikhila Ravi, Laurens van der Maaten, Armand Joulin, Ishan Misra

HIGHLIGHT: Instead, in this paper, we propose a single model which excels at classifying images, videos, and single-view 3D data using exactly the same model parameters.

949, TITLE: Neural Shape Mating: Self-Supervised Object Assembly With Adversarial Shape Priors

http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Neural_Shape_Mating_Self-Supervised_Object_Assembly_With_Adversarial_Shape_Priors_CVPR_2022_paper.html

AUTHORS: Yun-Chun Chen, Haoda Li, Dylan Turpin, Alec Jacobson, Animesh Garg

HIGHLIGHT: By focusing on shape alignment rather than semantic cues, we can achieve across category generalization and scaling. In this paper, we introduce a novel task, pairwise 3D geometric shape mating, and propose Neural Shape Mating (NSM) to tackle this problem.

950, TITLE: Reflash Dropout in Image Super-Resolution

http://openaccess.thecvf.com//content/CVPR2022/html/Kong_Reflash_Dropout_in_Image_Super-Resolution_CVPR_2022_paper.html

AUTHORS: Xiangtao Kong, Xina Liu, Jinjin Gu, Yu Qiao, Chao Dong

HIGHLIGHT:		However, in this paper, we show that appropriate usage of dropout benefits SR networks and improves the generalization ability.

951, TITLE:		WildNet: Learning Domain Generalized Semantic Segmentation From the Wild
http://openaccess.thecvf.com//content/CVPR2022/html/Lee_WildNet_Learning_Domain_Generalized_Semantic_Segmentation_From_the_Wild_CVPR_2022_paper.html
AUTHORS:		Suhyeon Lee, Hongje Seong, Seongwon Lee, Euntai Kim
HIGHLIGHT:		In this paper, we propose to diversify both the content and style of the source domain with the help of the wild.

952, TITLE:		Auditing Privacy Defenses in Federated Learning via Generative Gradient Leakage
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Auditing_Privacy_Defenses_in_Federated_Learning_via_Generative_Gradient_Leakage_CVPR_2022_paper.html
AUTHORS:		Zhuohang Li, Jiaxin Zhang, Luyang Liu, Jian Liu
HIGHLIGHT:		In this work, we validate that the private training data can still be leaked under certain defense settings with a new type of leakage, i.e., Generative Gradient Leakage (GGL).

953, TITLE:		DAIR-V2X: A Large-Scale Dataset for Vehicle-Infrastructure Cooperative 3D Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Yu_DAIR-V2X_A_Large-Scale_Dataset_for_Vehicle-Infrastructure_Cooperative_3D_Object_Detection_CVPR_2022_paper.html
AUTHORS:		Haibao Yu, Yizhen Luo, Mao Shu, Yiyi Huo, Zebang Yang, Yifeng Shi, Zhenglong Guo, Hanyu Li, Xing Hu, Jirui Yuan, Zaiqing Nie
HIGHLIGHT:		DAIR-V2X comprises 71254 LiDAR frames and 71254 Camera frames, and all frames are captured from real scenes with 3D annotations.

954, TITLE:		DECORE: Deep Compression With Reinforcement Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Alwani_DECORE_Deep_Compression_With_Reinforcement_Learning_CVPR_2022_paper.html
AUTHORS:		Manoj Alwani, Yang Wang, Vashisht Madhavan
HIGHLIGHT:		We present DECORE, a reinforcement learning based approach to automate the network compression process.

955, TITLE:		Time3D: End-to-End Joint Monocular 3D Object Detection and Tracking for Autonomous Driving
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Time3D_End-to-End_Joint_Monocular_3D_Object_Detection_and_Tracking_for_CVPR_2022_paper.html
AUTHORS:		Peixuan Li, Jieyu Jin
HIGHLIGHT:		In this work, we propose jointly training 3D detection and 3D tracking from only monocular videos in an end-to-end manner.

956, TITLE:		MonoJSG: Joint Semantic and Geometric Cost Volume for Monocular 3D Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Lian_MonoJSG_Joint_Semantic_and_Geometric_Cost_Volume_for_Monocular_3D_CVPR_2022_paper.html
AUTHORS:		Qing Lian, Peiliang Li, Xiaozhi Chen
HIGHLIGHT:		To benefit from both the powerful feature representation in DNN and pixel-level geometric constraints, we reformulate the monocular object depth estimation as a progressive refinement problem and propose a joint semantic and geometric cost volume to model the depth error.

957, TITLE:		Task Discrepancy Maximization for Fine-Grained Few-Shot Classification
http://openaccess.thecvf.com//content/CVPR2022/html/Lee_Task_Discrepancy_Maximization_for_Fine-Grained_Few-Shot_Classification_CVPR_2022_paper.html
AUTHORS:		SuBeen Lee, WonJun Moon, Jae-Pil Heo
HIGHLIGHT:		Our objective is to localize the class-wise discriminative regions by highlighting channels encoding distinct information of the class.

958, TITLE:		FedDC: Federated Learning With Non-IID Data via Local Drift Decoupling and Correction
http://openaccess.thecvf.com//content/CVPR2022/html/Gao_FedDC_Federated_Learning_With_Non-IID_Data_via_Local_Drift_Decoupling_CVPR_2022_paper.html
AUTHORS:		Liang Gao, Huazhu Fu, Li Li, Yingwen Chen, Ming Xu, Cheng-Zhong Xu
HIGHLIGHT:		However, the key challenge in federated learning is that the clients have significant statistical heterogeneity among their local data distributions, which would cause inconsistent optimized local models on the client-side. To address this fundamental dilemma, we propose a novel federated learning algorithm with local drift decoupling and correction (FedDC).

959, TITLE:		Efficient Classification of Very Large Images With Tiny Objects
http://openaccess.thecvf.com//content/CVPR2022/html/Kong_Efficient_Classification_of_Very_Large_Images_With_Tiny_Objects_CVPR_2022_paper.html
AUTHORS:		Fanjie Kong, Ricardo Henao
HIGHLIGHT:		We present an end-to-end CNN model termed Zoom-In network that leverages hierarchical attention sampling for classification of large images with tiny objects using a single GPU.

960, TITLE:		SWEM: Towards Real-Time Video Object Segmentation With Sequential Weighted Expectation-Maximization
http://openaccess.thecvf.com//content/CVPR2022/html/Lin_SWEM_Towards_Real-Time_Video_Object_Segmentation_With_Sequential_Weighted_Expectation-Maximization_CVPR_2022_paper.html

AUTHORS: Zhihui Lin, Tianyu Yang, Maomao Li, Ziyu Wang, Chun Yuan, Wenhao Jiang, Wei Liu
HIGHLIGHT: However, continuously growing and redundant template features lead to an inefficient inference. To alleviate this, we propose a novel Sequential Weighted Expectation-Maximization (SWEM) network to greatly reduce the redundancy of memory features.

961, TITLE: Point-to-Voxel Knowledge Distillation for LiDAR Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Hou_Point-to-Voxel_Knowledge_Distillation_for_LiDAR_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Yuenan Hou, Xinge Zhu, Yuexin Ma, Chen Change Loy, Yikang Li
HIGHLIGHT: To tackle the aforementioned problems, we propose the Point-to-Voxel Knowledge Distillation (PVD), which transfers the hidden knowledge from both point level and voxel level.

962, TITLE: Leveling Down in Computer Vision: Pareto Inefficiencies in Fair Deep Classifiers
http://openaccess.thecvf.com//content/CVPR2022/html/Zietlow_Leveling_Down_in_Computer_Vision_Pareto_Inefficiencies_in_Fair_Deep_CVPR_2022_paper.html
AUTHORS: Dominik Zietlow, Michael Lohaus, Guha Balakrishnan, Matthäus Kleindessner, Francesco Locatello, Bernhard Schölkopf, Chris Russell
HIGHLIGHT: We corroborate this analysis with extensive experimental support that shows that many of the fairness heuristics used in computer vision also degrade performance on the most disadvantaged groups. Building on these insights, we propose an adaptive augmentation strategy that, uniquely, of all methods tested, improves performance for the disadvantaged groups.

963, TITLE: Generating Diverse 3D Reconstructions From a Single Occluded Face Image
http://openaccess.thecvf.com//content/CVPR2022/html/Dey_Generating_Diverse_3D_Reconstructions_From_a_Single_Occluded_Face_Image_CVPR_2022_paper.html
AUTHORS: Rahul Dey, Vishnu Naresh Boddeti
HIGHLIGHT: Furthermore, while a plurality of 3D reconstructions is plausible in the occluded regions, existing approaches are limited to generating only a single solution. To address both of these challenges, we present Diverse3DFace, which is specifically designed to simultaneously generate a diverse and realistic set of 3D reconstructions from a single occluded face image.

964, TITLE: RBGNet: Ray-Based Grouping for 3D Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_RBGNet_Ray-Based_Grouping_for_3D_Object_Detection_CVPR_2022_paper.html
AUTHORS: Haiyang Wang, Shaoshuai Shi, Ze Yang, Rongyao Fang, Qi Qian, Hongsheng Li, Bernt Schiele, Liwei Wang
HIGHLIGHT: In this paper, we propose the RBGNet framework, a voting-based 3D detector for accurate 3D object detection from point clouds.

965, TITLE: Stand-Alone Inter-Frame Attention in Video Models
http://openaccess.thecvf.com//content/CVPR2022/html/Long_Stand-Alone_Inter-Frame_Attention_in_Video_Models_CVPR_2022_paper.html
AUTHORS: Fuchen Long, Zhaofan Qiu, Yingwei Pan, Ting Yao, Jiebo Luo, Tao Mei
HIGHLIGHT: In this paper, we present a new recipe of inter-frame attention block, namely Stand-alone Inter-Frame Attention (SIFA), that novelly delves into the deformation across frames to estimate local self-attention on each spatial location.

966, TITLE: Uncertainty-Aware Adaptation for Self-Supervised 3D Human Pose Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Kundu_Uncertainty-Aware_Adaptation_for_Self-Supervised_3D_Human_Pose_Estimation_CVPR_2022_paper.html
AUTHORS: Jogendra Nath Kundu, Siddharth Seth, Pradyumna YM, Varun Jampani, Anirban Chakraborty, R. Venkatesh Babu
HIGHLIGHT: We introduce MRP-Net that constitutes a common deep network backbone with two output heads subscribing to two diverse configurations; a) model-free joint localization and b) model-based parametric regression.

967, TITLE: Open-Domain, Content-Based, Multi-Modal Fact-Checking of Out-of-Context Images via Online Resources
http://openaccess.thecvf.com//content/CVPR2022/html/Abdelnabi_Open-Domain_Content-Based_Multi-Modal_Fact-Checking_of_Out-of-Context_Images_via_Online_Resources_CVPR_2022_paper.html
AUTHORS: Sahar Abdelnabi, Rakibul Hasan, Mario Fritz
HIGHLIGHT: The internet is being used as the go-to way to verify information using different sources and modalities. Our goal is an inspectable method that automates this time-consuming and reasoning-intensive process by fact-checking the image-caption pairing using Web evidence.

968, TITLE: Memory-Augmented Deep Conditional Unfolding Network for Pan-Sharpening
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Memory-Augmented_Deep_Conditional_Unfolding_Network_for_Pan-Sharpening_CVPR_2022_paper.html
AUTHORS: Gang Yang, Man Zhou, Keyu Yan, Aiping Liu, Xueyang Fu, Fan Wang
HIGHLIGHT: Additionally, they ignore the different characteristics of each band of MS images and directly concatenate them with panchromatic (PAN) images, leading to severe copy artifacts. To address the above issues, we propose an interpretable deep neural network, namely Memory-augmented Deep Conditional Unfolding Network with two specified core designs.

969, TITLE: Semi-Supervised Wide-Angle Portraits Correction by Multi-Scale Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Zhu_Semi-Supervised_Wide-Angle_Portraits_Correction_by_Multi-Scale_Transformer_CVPR_2022_paper.html

AUTHORS: Fushun Zhu, Shan Zhao, Peng Wang, Hao Wang, Hua Yan, Shuaicheng Liu
HIGHLIGHT: In this work, we design a semi-supervised scheme and build a high-quality unlabeled dataset with rich scenarios, allowing us to simultaneously use labeled and unlabeled data to improve performance.

970, TITLE: Large-Scale Pre-Training for Person Re-Identification With Noisy Labels
http://openaccess.thecvf.com//content/CVPR2022/html/Fu_Large-Scale_Pre-Training_for_Person_Re-Identification_With_Noisy_Labels_CVPR_2022_paper.html
AUTHORS: Dengpan Fu, Dongdong Chen, Hao Yang, Jianmin Bao, Lu Yuan, Lei Zhang, Houqiang Li, Fang Wen, Dong Chen
HIGHLIGHT: This paper aims to address the problem of pre-training for person re-identification (Re-ID) with noisy labels.

971, TITLE: Adiabatic Quantum Computing for Multi Object Tracking
http://openaccess.thecvf.com//content/CVPR2022/html/Zaech_Adiabatic_Quantum_Computing_for_Multi_Object_Tracking_CVPR_2022_paper.html
AUTHORS: Jan-Nico Zaech, Alexander Liniger, Martin Danelljan, Dengxin Dai, Luc Van Gool
HIGHLIGHT: However, current MOT formulations are unsuitable for quantum computing due to their scaling properties. In this work, we therefore propose the first MOT formulation designed to be solved with AQC.

972, TITLE: Feature Erasing and Diffusion Network for Occluded Person Re-Identification
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Feature_Erasing_and_Diffusion_Network_for_Occluded_Person_Re-Identification_CVPR_2022_paper.html
AUTHORS: Zhikang Wang, Feng Zhu, Shixiang Tang, Rui Zhao, Lihuo He, Jiangning Song
HIGHLIGHT: In this paper, we propose a novel Feature Erasing and Diffusion Network (FED) to simultaneously handle challenges from NPO and NTP.

973, TITLE: Is Mapping Necessary for Realistic PointGoal Navigation?
http://openaccess.thecvf.com//content/CVPR2022/html/Partsey_Is_Mapping_Necessary_for_Realistic_PointGoal_Navigation_CVPR_2022_paper.html
AUTHORS: Ruslan Partsey, Erik Wijmans, Naoki Yokoyama, Oles Dobosevych, Dhruv Batra, Oleksandr Maksymets
HIGHLIGHT: However, for PointNav in a realistic setting (RGB-D and actuation noise, no GPS+Compass), this is an open question; one we tackle in this paper.

974, TITLE: Node-Aligned Graph Convolutional Network for Whole-Slide Image Representation and Classification
http://openaccess.thecvf.com//content/CVPR2022/html/Guan_Node-Aligned_Graph_Convolutional_Network_for_Whole-Slide_Image_Representation_and_Classification_CVPR_2022_paper.html
AUTHORS: Yonghang Guan, Jun Zhang, Kuan Tian, Sen Yang, Pei Dong, Jinxi Xiang, Wei Yang, Junzhou Huang, Yuyao Zhang, Xiao Han
HIGHLIGHT: In this paper, we propose a hierarchical global-to-local clustering strategy to build a Node-Aligned GCN (NAGCN) to represent WSI with rich local structural information as well as global distribution.

975, TITLE: Represent, Compare, and Learn: A Similarity-Aware Framework for Class-Agnostic Counting
http://openaccess.thecvf.com//content/CVPR2022/html/Shi_Represent_Compare_and_Learn_A_Similarity-Aware_Framework_for_Class-Agnostic_Counting_CVPR_2022_paper.html
AUTHORS: Min Shi, Hao Lu, Chen Feng, Chengxin Liu, Zhiguo Cao
HIGHLIGHT: In this work, we propose a similarity-aware CAC framework that jointly learns representation and similarity metric.

976, TITLE: Masked Feature Prediction for Self-Supervised Visual Pre-Training
http://openaccess.thecvf.com//content/CVPR2022/html/Wei_Masked_Feature_Prediction_for_Self-Supervised_Visual_Pre-Training_CVPR_2022_paper.html
AUTHORS: Chen Wei, Haoqi Fan, Saining Xie, Chao-Yuan Wu, Alan Yuille, Christoph Feichtenhofer
HIGHLIGHT: We present Masked Feature Prediction (MaskFeat) for self-supervised pre-training of video models.

977, TITLE: Critical Regularizations for Neural Surface Reconstruction in the Wild
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Critical_Regularizations_for_Neural_Surface_Reconstruction_in_the_Wild_CVPR_2022_paper.html
AUTHORS: Jingyang Zhang, Yao Yao, Shiwei Li, Tian Fang, David McKinnon, Yanghai Tsin, Long Quan
HIGHLIGHT: In this paper, we present RegSDF, which shows that proper point cloud supervisions and geometry regularizations are sufficient to produce high-quality and robust reconstruction results.

978, TITLE: EASE: Unsupervised Discriminant Subspace Learning for Transductive Few-Shot Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Zhu_EASE_Unsupervised_Discriminant_Subspace_Learning_for_Transductive_Few-Shot_Learning_CVPR_2022_paper.html
AUTHORS: Hao Zhu, Piotr Koniusz
HIGHLIGHT: We present an unsupErvised discriminAnt Subspace lEarning (EASE) that improves transductive few-shot learning performance by learning a linear projection onto a subspace built from features of the support set and the unlabeled query set in the test time.

979, TITLE: Object-Relation Reasoning Graph for Action Recognition

http://openaccess.thecvf.com//content/CVPR2022/html/Ou_Object-Relation_Reasoning_Graph_for_Action_Recognition_CVPR_2022_paper.html
AUTHORS: Yangjun Ou, Li Mi, Zhenzhong Chen
HIGHLIGHT: In this paper, we propose an Object-Relation Reasoning Graph (OR2G) for reasoning about action in videos.

980, TITLE: Semantic Segmentation by Early Region Proxy
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Semantic_Segmentation_by_Early_Region_Proxy_CVPR_2022_paper.html
AUTHORS: Yifan Zhang, Bo Pang, Cewu Lu
HIGHLIGHT: In this work, we present a novel and efficient modeling that starts from interpreting the image as a tessellation of learnable regions, each of which has flexible geometrics and carries homogeneous semantics.

981, TITLE: GIQE: Generic Image Quality Enhancement via Nth Order Iterative Degradation
http://openaccess.thecvf.com//content/CVPR2022/html/Shyam_GIQE_Generic_Image_Quality_Enhancement_via_Nth_Order_Iterative_Degradation_CVPR_2022_paper.html
AUTHORS: Pranjay Shyam, Kyung-Soo Kim, Kuk-Jin Yoon
HIGHLIGHT: In this paper, we utilize a synthetic degradation model that recursively applies sets of random degradations to generate naturalistic degradation images of varying complexity, which are used as input.

982, TITLE: Instance Segmentation With Mask-Supervised Polygonal Boundary Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Lazarow_Instance_Segmentation_With_Mask-Supervised_Polygonal_Boundary_Transformers_CVPR_2022_paper.html
AUTHORS: Justin Lazarow, Weijian Xu, Zhuowen Tu
HIGHLIGHT: In this paper, we present an end-to-end instance segmentation method that regresses a polygonal boundary for each object instance.

983, TITLE: FaceVerse: A Fine-Grained and Detail-Controllable 3D Face Morphable Model From a Hybrid Dataset
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_FaceVerse_A_Fine-Grained_and_Detail-Controllable_3D_Face_Morphable_Model_From_CVPR_2022_paper.html
AUTHORS: Lizhen Wang, Zhiyuan Chen, Tao Yu, Chenguang Ma, Liang Li, Yebin Liu
HIGHLIGHT: We present FaceVerse, a fine-grained 3D Neural Face Model, which is built from hybrid East Asian face datasets containing 60K fused RGB-D images and 2K high-fidelity 3D head scan models.

984, TITLE: Bring Evanescent Representations to Life in Lifelong Class Incremental Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Toldo_Bring_Evanescent_Representations_to_Life_in_Lifelong_Class_Incremental_Learning_CVPR_2022_paper.html
AUTHORS: Marco Toldo, Mete Ozay
HIGHLIGHT: However, if not updated, those representations become increasingly outdated as the incremental learning progresses with new classes. To address the aforementioned problems, we propose a framework which aims to (i) model the semantic drift by learning the relationship between representations of past and novel classes among incremental steps, and (ii) estimate the feature drift, defined as the evolution of the representations learned by models at each incremental step.

985, TITLE: Single-Stage 3D Geometry-Preserving Depth Estimation Model Training on Dataset Mixtures With Uncalibrated Stereo Data
http://openaccess.thecvf.com//content/CVPR2022/html/Patakin_Single-Stage_3D_Geometry-Preserving_Depth_Estimation_Model_Training_on_Dataset_Mixtures_CVPR_2022_paper.html
AUTHORS: Nikolay Patakin, Anna Vorontsova, Mikhail Artemyev, Anton Konushin
HIGHLIGHT: On the contrary, we propose GP2, General-Purpose and Geometry-Preserving training scheme, and show that conventional SVDE models can learn correct shifts themselves without any post-processing, benefiting from using stereo data even in the geometry-preserving setting.

986, TITLE: LD-ConGR: A Large RGB-D Video Dataset for Long-Distance Continuous Gesture Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_LD-ConGR_A_Large_RGB-D_Video_Dataset_for_Long-Distance_Continuous_Gesture_CVPR_2022_paper.html
AUTHORS: Dan Liu, Libo Zhang, Yanjun Wu
HIGHLIGHT: To apply gesture recognition to long-distance interactive scenes such as meetings and smart homes, a large RGB-D video dataset LD-ConGR is established in this paper.

987, TITLE: SimVQA: Exploring Simulated Environments for Visual Question Answering
http://openaccess.thecvf.com//content/CVPR2022/html/Cascante-Bonilla_SimVQA_Exploring_Simulated_Environments_for_Visual_Question_Answering_CVPR_2022_paper.html
AUTHORS: Paola Cascante-Bonilla, Hui Wu, Letao Wang, Rogerio S. Feris, Vicente Ordonez
HIGHLIGHT: In this work we explore using synthetic computer-generated data to fully control the visual and language space, allowing us to provide more diverse scenarios.

988, TITLE: Thin-Plate Spline Motion Model for Image Animation
http://openaccess.thecvf.com//content/CVPR2022/html/Zhao_Thin-Plate_Spline_Motion_Model_for_Image_Animation_CVPR_2022_paper.html
AUTHORS: Jian Zhao, Hui Zhang

HIGHLIGHT: However, it remains a significant challenge for current unsupervised methods when there is a large pose gap between the objects in the source and driving images. In this paper, a new end-to-end unsupervised motion transfer framework is proposed to overcome such issue.

989, TITLE: Learning Local Displacements for Point Cloud Completion
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Learning_Local_Displacements_for_Point_Cloud_Completion_CVPR _2022_paper.html
AUTHORS: Yida Wang, David Joseph Tan, Nassir Navab, Federico Tombari
HIGHLIGHT: We propose a novel approach aimed at object and semantic scene completion from a partial scan represented as a 3D point cloud.

990, TITLE: Human Hands As Probes for Interactive Object Understanding
http://openaccess.thecvf.com//content/CVPR2022/html/Goyal_Human_Hands_As_Probes_for_Interactive_Object_Understanding_C VPR_2022_paper.html
AUTHORS: Mohit Goyal, Sahil Modi, Rishabh Goyal, Saurabh Gupta
HIGHLIGHT: Interactive object understanding, or what we can do to objects and how is a long-standing goal of computer vision. In this paper, we tackle this problem through observation of human hands in in-the-wild egocentric videos.

991, TITLE: Understanding and Increasing Efficiency of Frank-Wolfe Adversarial Training
http://openaccess.thecvf.com//content/CVPR2022/html/Tsiligkaridis_Understanding_and_Increasing_Efficiency_of_Frank-Wolfe_Adversarial_Training_CVPR_2022_paper.html
AUTHORS: Theodoros Tsiligkaridis, Jay Roberts
HIGHLIGHT: We develop a theoretical framework for adversarial training with FW optimization (FW-AT) that reveals a geometric connection between the loss landscape and the distortion of l-inf FW attacks (the attack's l-2 norm).

992, TITLE: Certified Patch Robustness via Smoothed Vision Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Salman_Certified_Patch_Robustness_via_Smoothed_Vision_Transformers_C VPR_2022_paper.html
AUTHORS: Hadi Salman, Saachi Jain, Eric Wong, Aleksander Madry
HIGHLIGHT: We demonstrate how using vision transformers enables significantly better certified patch robustness that is also more computationally efficient and does not incur a substantial drop in standard accuracy.

993, TITLE: Look Back and Forth: Video Super-Resolution With Explicit Temporal Difference Modeling
http://openaccess.thecvf.com//content/CVPR2022/html/Isobe_Look_Back_and_Forth_Video_Super-Resolution_With_Explicit_Temporal_Difference_CVPR_2022_paper.html
AUTHORS: Takashi Isobe, Xu Jia, Xin Tao, Changlin Li, Ruihuang Li, Yongjie Shi, Jing Mu, Huchuan Lu, Yu-Wing Tai
HIGHLIGHT: In this paper, we propose to explore the role of explicit temporal difference modeling in both LR and HR space.

994, TITLE: UCC: Uncertainty Guided Cross-Head Co-Training for Semi-Supervised Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Fan_UCC_Uncertainty_Guided_Cross-Head_Co-Training_for_Semi-Supervised_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Jiashuo Fan, Bin Gao, Huan Jin, Lihui Jiang
HIGHLIGHT: We present a novel learning framework called Uncertainty guided Cross-head Co-training (UCC) for semi-supervised semantic segmentation.

995, TITLE: HVH: Learning a Hybrid Neural Volumetric Representation for Dynamic Hair Performance Capture
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_HVH_Learning_a_Hybrid_Neural_Volumetric_Representation_for_Dy namic_Hair_CVPR_2022_paper.html
AUTHORS: Ziyan Wang, Giljoo Nam, Tuur Stuyck, Stephen Lombardi, Michael Zollhöfer, Jessica Hodgins, Christoph Lassner
HIGHLIGHT: In this paper, we address the aforementioned problems: 1) we use a novel, volumetric hair representation that is composed of thousands of primitives.

996, TITLE: RADU: Ray-Aligned Depth Update Convolutions for ToF Data Denoising
http://openaccess.thecvf.com//content/CVPR2022/html/Schelling_RADU_Ray-Aligned_Depth_Update_Convolutions_for_ToF_Data_Denoising_CVPR_2022_paper.html
AUTHORS: Michael Schelling, Pedro Hermosilla, Timo Ropinski
HIGHLIGHT: In this paper, we propose an iterative correcting approach operating in 3D space, that is designed to learn on 2.5D data by enabling 3D point convolutions to correct the points' positions along the view direction.

997, TITLE: Rethinking Visual Geo-Localization for Large-Scale Applications
http://openaccess.thecvf.com//content/CVPR2022/html/Berton_Rethinking_Visual_Geo-Localization_for_Large-Scale_Applications_CVPR_2022_paper.html
AUTHORS: Gabriele Berton, Carlo Masone, Barbara Caputo
HIGHLIGHT: We find that current methods fail to scale to such large datasets, therefore we design a new highly scalable training technique, called CosPlace, which casts the training as a classification problem avoiding the expensive mining needed by the commonly used contrastive learning.

998, TITLE: Learning Based Multi-Modality Image and Video Compression

http://openaccess.thecvf.com//content/CVPR2022/html/Lu_Learning_Based_Multi-Modality_Image_and_Video_Compression_CVPR_2022_paper.html
AUTHORS: Guo Lu, Tianxiong Zhong, Jing Geng, Qiang Hu, Dong Xu
HIGHLIGHT: This work proposes a multi-modality compression framework for infrared and visible image pairs by exploiting the cross-modality redundancy.

999, TITLE: A Stitch in Time Saves Nine: A Train-Time Regularizing Loss for Improved Neural Network Calibration
http://openaccess.thecvf.com//content/CVPR2022/html/Hebbalaguppe_A_Stitch_in_Time_Saves_Nine_A_Train-Time_Regularizing_Loss_CVPR_2022_paper.html
AUTHORS: Ramya Hebbalaguppe, Jatin Prakash, Neelabh Madan, Chetan Arora
HIGHLIGHT: In this paper, we argue for intervening at the train time itself, so as to directly produce calibrated DNN models.

1000, TITLE: The Principle of Diversity: Training Stronger Vision Transformers Calls for Reducing All Levels of Redundancy
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_The_Principle_of_Diversity_Training_Stronger_Vision_Transformers_Calls_for_CVPR_2022_paper.html
AUTHORS: Tianlong Chen, Zhenyu Zhang, Yu Cheng, Ahmed Awadallah, Zhangyang Wang
HIGHLIGHT: In view of them, we advocate a principle of diversity for training ViTs, by presenting corresponding regularizers that encourage the representation diversity and coverage at each of those levels, that enabling capturing more discriminative information.

1001, TITLE: Deep Image-Based Illumination Harmonization
http://openaccess.thecvf.com//content/CVPR2022/html/Bao_Deep_Image-Based_Illumination_Harmonization_CVPR_2022_paper.html
AUTHORS: Zhongyun Bao, Chengjiang Long, Gang Fu, Daquan Liu, Yuanzhen Li, Jiaming Wu, Chunxia Xiao
HIGHLIGHT: In this paper,we formulate seamless illumination harmonization as anillumination exchange and aggregation problem.

1002, TITLE: ViM: Out-of-Distribution With Virtual-Logit Matching
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_ViM_Out-of-Distribution_With_Virtual-Logit_Matching_CVPR_2022_paper.html
AUTHORS: Haoqi Wang, Zhizhong Li, Litong Feng, Wayne Zhang
HIGHLIGHT: There are OOD samples that are easy to identify in the feature space while hard to distinguish in the logit space and vice versa. Motivated by this observation, we propose a novel OOD scoring method named Virtual-logit Matching (ViM), which combines the class-agnostic score from feature space and the In-Distribution (ID) class-dependent logits.

1003, TITLE: Active Learning by Feature Mixing
http://openaccess.thecvf.com//content/CVPR2022/html/Parvaneh_Active_Learning_by_Feature_Mixing_CVPR_2022_paper.html
AUTHORS: Amin Parvaneh, Ehsan Abbasnejad, Damien Teney, Gholamreza (Reza) Haffari, Anton van den Hengel, Javen Qinfeng Shi
HIGHLIGHT: In this paper, we propose a novel method for batch AL called ALFA-Mix.

1004, TITLE: Towards Accurate Facial Landmark Detection via Cascaded Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Towards_Accurate_Facial_Landmark_Detection_via_Cascaded_Transformers_CVPR_2022_paper.html
AUTHORS: Hui Li, Zidong Guo, Seon-Min Rhee, Seungju Han, Jae-Joon Han
HIGHLIGHT: In this paper, an accurate facial landmark detector is proposed based on cascaded transformers.

1005, TITLE: Class-Aware Contrastive Semi-Supervised Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Class-Aware_Contrastive_Semi-Supervised_Learning_CVPR_2022_paper.html
AUTHORS: Fan Yang, Kai Wu, Shuyi Zhang, Guannan Jiang, Yong Liu, Feng Zheng, Wei Zhang, Chengjie Wang, Long Zeng
HIGHLIGHT: Moreover, the model's judgment becomes noisier in real-world applications with extensive out-of-distribution data. To address this issue, we propose a general method named Class-aware Contrastive Semi-Supervised Learning (CCSSL), which is a drop-in helper to improve the pseudo-label quality and enhance the model's robustness in the real-world setting.

1006, TITLE: Long-Term Visual Map Sparsification With Heterogeneous GNN
http://openaccess.thecvf.com//content/CVPR2022/html/Chang_Long-Term_Visual_Map_Sparsification_With_Heterogeneous_GNN_CVPR_2022_paper.html
AUTHORS: Ming-Fang Chang, Yipu Zhao, Rajvi Shah, Jakob J. Engel, Michael Kaess, Simon Lucey
HIGHLIGHT: In this paper, we aim to overcome the environmental changes and reduce the map size at the same time by selecting points that are valuable to future localization.

1007, TITLE: Debiased Learning From Naturally Imbalanced Pseudo-Labels
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Debiased_Learning_From_Naturally_Imbalanced_Pseudo-Labels_CVPR_2022_paper.html
AUTHORS: Xudong Wang, Zhirong Wu, Long Lian, Stella X. Yu
HIGHLIGHT: To eliminate the model bias, we propose a simple yet effective method DebiasMatch, comprising of an adaptive debiasing module and an adaptive marginal loss.

1008, TITLE: RNNPose: Recurrent 6-DoF Object Pose Refinement With Robust Correspondence Field Estimation and Pose Optimization
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_RNNPose_Recurrent_6-DoF_Object_Pose_Refinement_With_Robust_Correspondence_Field_CVPR_2022_paper.html
AUTHORS: Yan Xu, Kwan-Yee Lin, Guofeng Zhang, Xiaogang Wang, Hongsheng Li
HIGHLIGHT: In this paper, we propose a framework based on a recurrent neural network (RNN) for object pose refinement, which is robust to erroneous initial poses and occlusions.

1009, TITLE: Ditto: Building Digital Twins of Articulated Objects From Interaction
http://openaccess.thecvf.com//content/CVPR2022/html/Jiang_Ditto_Building_Digital_Twins_of_Articulated_Objects_From_Interaction_CVPR_2022_paper.html
AUTHORS: Zhenyu Jiang, Cheng-Chun Hsu, Yuke Zhu
HIGHLIGHT: We introduce Ditto to learn articulation model estimation and 3D geometry reconstruction of an articulated object through interactive perception.

1010, TITLE: Dual-AI: Dual-Path Actor Interaction Learning for Group Activity Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Han_Dual-AI_Dual-Path_Actor_Interaction_Learning_for_Group_Activity_Recognition_CVPR_2022_paper.html
AUTHORS: Mingfei Han, David Junhao Zhang, Yali Wang, Rui Yan, Lina Yao, Xiaojun Chang, Yu Qiao
HIGHLIGHT: Hence, it is often difficult to model complex group activities from a single view of spatial-temporal actor evolution. To tackle this problem, we propose a distinct Dual-path Actor Interaction (Dual-AI) framework, which flexibly arranges spatial and temporal transformers in two complementary orders, enhancing actor relations by integrating merits from different spatio-temporal paths.

1011, TITLE: Harmony: A Generic Unsupervised Approach for Disentangling Semantic Content From Parameterized Transformations
http://openaccess.thecvf.com//content/CVPR2022/html/Uddin_Harmony_A_Generic_Unsupervised_Approach_for_Disentangling_Semantic_Content_From_CVPR_2022_paper.html
AUTHORS: Mostofa Rafid Uddin, Gregory Howe, Xiangrui Zeng, Min Xu
HIGHLIGHT: Disentangling images into semantic content factors and transformations can provide significant benefits into many domain-specific image analysis tasks. To this end, we propose a generic unsupervised framework, Harmony, that simultaneously and explicitly disentangles semantic content from multiple parameterized transformations.

1012, TITLE: Talking Face Generation With Multilingual TTS
http://openaccess.thecvf.com//content/CVPR2022/html/Song_Talking_Face_Generation_With_Multilingual_TTS_CVPR_2022_paper.html
AUTHORS: Hyoung-Kyu Song, Sang Hoon Woo, Junhyeok Lee, Seungmin Yang, Hyunjae Cho, Youseong Lee, Dongho Choi, Kang-wook Kim
HIGHLIGHT: In this work, we introduce a talking face generation system that generalizes to different languages.

1013, TITLE: A Brand New Dance Partner: Music-Conditioned Pluralistic Dancing Controlled by Multiple Dance Genres
http://openaccess.thecvf.com//content/CVPR2022/html/Kim_A_Brand_New_Dance_Partner_Music-Conditioned_Pluralistic_Dancing_Controlled_by_CVPR_2022_paper.html
AUTHORS: Jinwoo Kim, Heeseok Oh, Seongjean Kim, Hoseok Tong, Sanghoon Lee
HIGHLIGHT: Yet, a serious limitation remains that all existing algorithms can return repeated patterns for a given initial pose sequence, which may be inferior. To mitigate this issue, we propose MNET, a novel and scalable approach that can perform music-conditioned pluralistic dance generation synthesized by multiple dance genres using only a single model.

1014, TITLE: Kernelized Few-Shot Object Detection With Efficient Integral Aggregation
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Kernelized_Few-Shot_Object_Detection_With_Efficient_Integral_Aggregation_CVPR_2022_paper.html
AUTHORS: Shan Zhang, Lei Wang, Naila Murray, Piotr Koniusz
HIGHLIGHT: We design a Kernelized Few-shot Object Detector by leveraging kernelized matrices computed over multiple proposal regions, which yield expressive non-linear representations whose model complexity is learned on the fly.

1015, TITLE: Transformer Based Line Segment Classifier With Image Context for Real-Time Vanishing Point Detection in Manhattan World
http://openaccess.thecvf.com//content/CVPR2022/html/Tong_Transformer_Based_Line_Segment_Classifier_With_Image_Context_for_Real-Time_CVPR_2022_paper.html
AUTHORS: Xin Tong, Xianghua Ying, Yongjie Shi, Ruibin Wang, Jinfa Yang
HIGHLIGHT: We find that image context can also contribute to accurate line classification. Based on this observation, we propose to classify line segments into three groups according to three unknown-but-sought vanishing points with Manhattan world assumption, using both geometric information and image context in this work.

1016, TITLE: Self-Sustaining Representation Expansion for Non-Exemplar Class-Incremental Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Zhu_Self-Sustaining_Representation_Expansion_for_Non-Exemplar_Class-Incremental_Learning_CVPR_2022_paper.html
AUTHORS: Kai Zhu, Wei Zhai, Yang Cao, Jiebo Luo, Zheng-Jun Zha

HIGHLIGHT: It is a challenging task since representation optimization and feature retention can only be achieved under supervision from new classes. To address this problem, we propose a novel self-sustaining representation expansion scheme.

1017, TITLE: Adaptive Early-Learning Correction for Segmentation From Noisy Annotations
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Adaptive_Early-Learning_Correction_for_Segmentation_From_Noisy_Annotations_CVPR_2022_paper.html
AUTHORS: Sheng Liu, Kangning Liu, Weicheng Zhu, Yiqiu Shen, Carlos Fernandez-Granda
HIGHLIGHT: In this work, we study the learning dynamics of deep segmentation networks trained on inaccurately-annotated data.

1018, TITLE: Cross-Domain Correlation Distillation for Unsupervised Domain Adaptation in Nighttime Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Gao_Cross-Domain_Correlation_Distillation_for_Unsupervised_Domain_Adaptation_in_Nighttime_Semantic_CVPR_2022_paper.html
AUTHORS: Huan Gao, Jichang Guo, Guoli Wang, Qian Zhang
HIGHLIGHT: Existing works, e.g., using the twilight as the intermediate target domain to perform the adaptation from daytime to nighttime, may fail to cope with the inherent difference between datasets caused by the camera equipment and the urban style. Faced with these two types of domain shifts, i.e., the illumination and the inherent difference of the datasets, we propose a novel domain adaptation framework via cross-domain correlation distillation, called CCDistill.

1019, TITLE: Context-Aware Video Reconstruction for Rolling Shutter Cameras
http://openaccess.thecvf.com//content/CVPR2022/html/Fan_Context-Aware_Video_Reconstruction_for_Rolling_Shutter_Cameras_CVPR_2022_paper.html
AUTHORS: Bin Fan, Yuchao Dai, Zhiyuan Zhang, Qi Liu, Mingyi He
HIGHLIGHT: However, these methods generate intermediate GS frames through image warping based on the RS model, which inevitably result in black holes and noticeable motion artifacts. In this paper, we alleviate these issues by proposing a context-aware GS video reconstruction architecture.

1020, TITLE: Towards Efficient Data Free Black-Box Adversarial Attack
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Towards_Efficient_Data_Free_Black-Box_Adversarial_Attack_CVPR_2022_paper.html
AUTHORS: Jie Zhang, Bo Li, Jianghe Xu, Shuang Wu, Shouhong Ding, Lei Zhang, Chao Wu
HIGHLIGHT: In this paper, by rethinking the collaborative relationship between the generator and the substitute model, we design a novel black-box attack framework.

1021, TITLE: Robust Contrastive Learning Against Noisy Views
http://openaccess.thecvf.com//content/CVPR2022/html/Chuang_Robust_Contrastive_Learning_Against_Noisy_Views_CVPR_2022_paper.html
AUTHORS: Ching-Yao Chuang, R Devon Hjelm, Xin Wang, Vibhav Vineet, Neel Joshi, Antonio Torralba, Stefanie Jegelka, Yale Song
HIGHLIGHT: In this work, we propose a new contrastive loss function that is robust against noisy views.

1022, TITLE: More Than Words: In-the-Wild Visually-Driven Prosody for Text-to-Speech
http://openaccess.thecvf.com//content/CVPR2022/html/Hassid_More_Than_Words_In-the-Wild_Visually-Driven_Prosody_for_Text-to-Speech_CVPR_2022_paper.html
AUTHORS: Michael Hassid, Michelle Tadmor Ramanovich, Brendan Shillingford, Miaosen Wang, Ye Jia, Tal Remez
HIGHLIGHT: In this paper we present VDTTS, a Visually-Driven Text-to-Speech model.

1023, TITLE: Cross-Modal Perceptionist: Can Face Geometry Be Gleaned From Voices?
http://openaccess.thecvf.com//content/CVPR2022/html/Wu_Cross-Modal_Perceptionist_Can_Face_Geometry_Be_Gleaned_From_Voices_CVPR_2022_paper.html
AUTHORS: Cho-Ying Wu, Chin-Cheng Hsu, Ulrich Neumann
HIGHLIGHT: We propose our analysis framework, Cross-Modal Perceptionist, under both supervised and unsupervised learning.

1024, TITLE: On Generalizing Beyond Domains in Cross-Domain Continual Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Simon_On_Generalizing_Beyond_Domains_in_Cross-Domain_Continual_Learning_CVPR_2022_paper.html
AUTHORS: Christian Simon, Masoud Faraki, Yi-Hsuan Tsai, Xiang Yu, Samuel Schulter, Yumin Suh, Mehrtash Harandi, Manmohan Chandraker
HIGHLIGHT: Many recent methods focus on preventing catastrophic forgetting under a typical assumption of thetrain and test data following a similar distribution. In thiswork, we consider the more realistic scenario of continuallearning under domain shifts where the model is able to gen-eralize its inference to a an unseen domain.

1025, TITLE: RSTT: Real-Time Spatial Temporal Transformer for Space-Time Video Super-Resolution
http://openaccess.thecvf.com//content/CVPR2022/html/Geng_RSTT_Real-Time_Spatial_Temporal_Transformer_for_Space-Time_Video_Super-Resolution_CVPR_2022_paper.html
AUTHORS: Zhicheng Geng, Luming Liang, Tianyu Ding, Ilya Zharkov

HIGHLIGHT:　　　　The existing methods based on Convolutional Neural Network (CNN) succeed in achieving visually satisfied results while suffer from slow inference speed due to their heavy architectures. We propose to resolve this issue by using a spatial-temporal transformer that naturally incorporates the spatial and temporal super resolution modules into a single model.

1026, TITLE:　　　　Learning Memory-Augmented Unidirectional Metrics for Cross-Modality Person Re-Identification
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Learning_Memory-Augmented_Unidirectional_Metrics_for_Cross-Modality_Person_Re-Identification_CVPR_2022_paper.html
AUTHORS:　　　　Jialun Liu, Yifan Sun, Feng Zhu, Hongbin Pei, Yi Yang, Wenhui Li
HIGHLIGHT:　　　　In response, we propose a Memory-Augmented Unidirectional Metric (MAUM) learning method consisting of two novel designs, i.e., unidirectional metrics, and memory-based augmentation.

1027, TITLE:　　　　A Closer Look at Few-Shot Image Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Zhao_A_Closer_Look_at_Few-Shot_Image_Generation_CVPR_2022_paper.html
AUTHORS:　　　　Yunqing Zhao, Henghui Ding, Houjing Huang, Ngai-Man Cheung
HIGHLIGHT:　　　　As our first contribution, we propose a framework to analyze existing methods during the adaptation.

1028, TITLE:　　　　Depth-Supervised NeRF: Fewer Views and Faster Training for Free
http://openaccess.thecvf.com//content/CVPR2022/html/Deng_Depth-Supervised_NeRF_Fewer_Views_and_Faster_Training_for_Free_CVPR_2022_paper.html
AUTHORS:　　　　Kangle Deng, Andrew Liu, Jun-Yan Zhu, Deva Ramanan
HIGHLIGHT:　　　　One potential reason is that standard volumetric rendering does not enforce the constraint that most of a scene's geometry consist of empty space and opaque surfaces. We formalize the above assumption through DS-NeRF (Depth-supervised Neural Radiance Fields), a loss for learning radiance fields that takes advantage of readily-available depth supervision.

1029, TITLE:　　　　Unsupervised Domain Generalization by Learning a Bridge Across Domains
http://openaccess.thecvf.com//content/CVPR2022/html/Harary_Unsupervised_Domain_Generalization_by_Learning_a_Bridge_Across_Domains_CVPR_2022_paper.html
AUTHORS:　　　　Sivan Harary, Eli Schwartz, Assaf Arbelle, Peter Staar, Shady Abu-Hussein, Elad Amrani, Roei Herzig, Amit Alfassy, Raja Giryes, Hilde Kuehne, Dina Katabi, Kate Saenko, Rogerio S. Feris, Leonid Karlinsky
HIGHLIGHT:　　　　In this paper, different from most cross-domain works that utilize some (or full) source domain supervision, we approach a relatively new and very practical Unsupervised Domain Generalization (UDG) setup of having no training supervision in neither source nor target domains.

1030, TITLE:　　　　Partial Class Activation Attention for Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Partial_Class_Activation_Attention_for_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS:　　　　Sun-Ao Liu, Hongtao Xie, Hai Xu, Yongdong Zhang, Qi Tian
HIGHLIGHT:　　　　Beyond the previous CAM generated from image-level classification, we present Partial CAM, which subdivides the task into region-level prediction and achieves better localization performance.

1031, TITLE:　　　　Multi-Scale Memory-Based Video Deblurring
http://openaccess.thecvf.com//content/CVPR2022/html/Ji_Multi-Scale_Memory-Based_Video_Deblurring_CVPR_2022_paper.html
AUTHORS:　　　　Bo Ji, Angela Yao
HIGHLIGHT:　　　　To achieve fine-grained deblurring, we designed a memory branch to memorize the blurry-sharp feature pairs in the memory bank, thus providing useful information for the blurry query input.

1032, TITLE:　　　　SkinningNet: Two-Stream Graph Convolutional Neural Network for Skinning Prediction of Synthetic Characters
http://openaccess.thecvf.com//content/CVPR2022/html/Mosella-Montoro_SkinningNet_Two-Stream_Graph_Convolutional_Neural_Network_for_Skinning_Prediction_of_CVPR_2022_paper.html
AUTHORS:　　　　Albert Mosella-Montoro, Javier Ruiz-Hidalgo
HIGHLIGHT:　　　　This work presents SkinningNet, an end-to-end Two-Stream Graph Neural Network architecture that computes skinning weights from an input mesh and its associated skeleton, without making any assumptions on shape class and structure of the provided mesh.

1033, TITLE:　　　　A Scalable Combinatorial Solver for Elastic Geometrically Consistent 3D Shape Matching
http://openaccess.thecvf.com//content/CVPR2022/html/Roetzer_A_Scalable_Combinatorial_Solver_for_Elastic_Geometrically_Consistent_3D_Shape_CVPR_2022_paper.html
AUTHORS:　　　　Paul Roetzer, Paul Swoboda, Daniel Cremers, Florian Bernard
HIGHLIGHT:　　　　We present a scalable combinatorial algorithm for globally optimizing over the space of geometrically consistent mappings between 3D shapes.

1034, TITLE:　　　　Learning Trajectory-Aware Transformer for Video Super-Resolution
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Learning_Trajectory-Aware_Transformer_for_Video_Super-Resolution_CVPR_2022_paper.html
AUTHORS:　　　　Chengxu Liu, Huan Yang, Jianlong Fu, Xueming Qian
HIGHLIGHT:　　　　We propose a novel Trajectory-aware Transformer for Video Super-Resolution (TTVSR).

1035, TITLE:　　　　Differentiable Dynamics for Articulated 3D Human Motion Reconstruction

http://openaccess.thecvf.com//content/CVPR2022/html/Gartner_Differentiable_Dynamics_for_Articulated_3D_Human_Motion_Reconstruction_CVPR_2022_paper.html
AUTHORS: Erik Gärtner, Mykhaylo Andriluka, Erwin Coumans, Cristian Sminchisescu
HIGHLIGHT: We introduce DiffPhy, a differentiable physics-based model for articulated 3d human motion reconstruction from video.

1036, TITLE: Geometric Structure Preserving Warp for Natural Image Stitching
http://openaccess.thecvf.com//content/CVPR2022/html/Du_Geometric_Structure_Preserving_Warp_for_Natural_Image_Stitching_CVPR_2022_paper.html
AUTHORS: Peng Du, Jifeng Ning, Jiguang Cui, Shaoli Huang, Xinchao Wang, Jiaxin Wang
HIGHLIGHT: However, most of the existing methods ignore the large-scale layouts reflected by straight lines or curves, decreasing overall stitching quality. To address this issue, this work presents a structure-preserving stitching approach that produces images with natural visual effects and less distortion.

1037, TITLE: GOAL: Generating 4D Whole-Body Motion for Hand-Object Grasping
http://openaccess.thecvf.com//content/CVPR2022/html/Taheri_GOAL_Generating_4D_Whole-Body_Motion_for_Hand-Object_Grasping_CVPR_2022_paper.html
AUTHORS: Omid Taheri, Vasileios Choutas, Michael J. Black, Dimitrios Tzionas
HIGHLIGHT: For the first time, we address the problem of generating full-body, hand and head motions of an avatar grasping an unknown object.

1038, TITLE: Multi-Robot Active Mapping via Neural Bipartite Graph Matching
http://openaccess.thecvf.com//content/CVPR2022/html/Ye_Multi-Robot_Active_Mapping_via_Neural_Bipartite_Graph_Matching_CVPR_2022_paper.html
AUTHORS: Kai Ye, Siyan Dong, Qingnan Fan, He Wang, Li Yi, Fei Xia, Jue Wang, Baoquan Chen
HIGHLIGHT: In this paper, we propose a novel algorithm, namely NeuralCoMapping, which takes advantage of both approaches.

1039, TITLE: Adversarial Texture for Fooling Person Detectors in the Physical World
http://openaccess.thecvf.com//content/CVPR2022/html/Hu_Adversarial_Texture_for_Fooling_Person_Detectors_in_the_Physical_World_CVPR_2022_paper.html
AUTHORS: Zhanhao Hu, Siyuan Huang, Xiaopei Zhu, Fuchun Sun, Bo Zhang, Xiaolin Hu
HIGHLIGHT: We propose a generative method, named Toroidal-Cropping-based Expandable Generative Attack (TC-EGA), to craft AdvTexture with repetitive structures.

1040, TITLE: Focal Length and Object Pose Estimation via Render and Compare
http://openaccess.thecvf.com//content/CVPR2022/html/Ponimatkin_Focal_Length_and_Object_Pose_Estimation_via_Render_and_Compare_CVPR_2022_paper.html
AUTHORS: Georgy Ponimatkin, Yann Labbé, Bryan Russell, Mathieu Aubry, Josef Sivic
HIGHLIGHT: We introduce FocalPose, a neural render-and-compare method for jointly estimating the camera-object 6D pose and camera focal length given a single RGB input image depicting a known object.

1041, TITLE: TO-FLOW: Efficient Continuous Normalizing Flows With Temporal Optimization Adjoint With Moving Speed
http://openaccess.thecvf.com//content/CVPR2022/html/Du_TO-FLOW_Efficient_Continuous_Normalizing_Flows_With_Temporal_Optimization_Adjoint_With_CVPR_2022_paper.html
AUTHORS: Shian Du, Yihong Luo, Wei Chen, Jian Xu, Delu Zeng
HIGHLIGHT: In this paper, a temporal optimization is proposed by optimizing the evolutionary time for forward propagation of the neural ODE training.

1042, TITLE: Arbitrary-Scale Image Synthesis
http://openaccess.thecvf.com//content/CVPR2022/html/Ntavelis_Arbitrary-Scale_Image_Synthesis_CVPR_2022_paper.html
AUTHORS: Evangelos Ntavelis, Mohamad Shahbazi, Iason Kastanis, Radu Timofte, Martin Danelljan, Luc Van Gool
HIGHLIGHT: We propose the design of scale-consistent positional encodings invariant to our generator's layers transformations.

1043, TITLE: Cross-Modal Representation Learning for Zero-Shot Action Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Lin_Cross-Modal_Representation_Learning_for_Zero-Shot_Action_Recognition_CVPR_2022_paper.html
AUTHORS: Chung-Ching Lin, Kevin Lin, Lijuan Wang, Zicheng Liu, Linjie Li
HIGHLIGHT: We present a cross-modal Transformer-based framework, which jointly encodes video data and text labels for zero-shot action recognition (ZSAR).

1044, TITLE: Conditional Prompt Learning for Vision-Language Models
http://openaccess.thecvf.com//content/CVPR2022/html/Zhou_Conditional_Prompt_Learning_for_Vision-Language_Models_CVPR_2022_paper.html
AUTHORS: Kaiyang Zhou, Jingkang Yang, Chen Change Loy, Ziwei Liu
HIGHLIGHT: To address the problem, we propose Conditional Context Optimization (CoCoOp), which extends CoOp by further learning a lightweight neural network to generate for each image an input-conditional token (vector).

1045, TITLE: Graph Sampling Based Deep Metric Learning for Generalizable Person Re-Identification

http://openaccess.thecvf.com//content/CVPR2022/html/Liao_Graph_Sampling_Based_Deep_Metric_Learning_for_Generalizable_Person_Re-Identification_CVPR_2022_paper.html
AUTHORS: Shengcai Liao, Ling Shao
HIGHLIGHT: To do so, in this paper, we propose an efficient mini-batch sampling method, called graph sampling (GS), for large-scale deep metric learning.

1046, TITLE: Retrieval-Based Spatially Adaptive Normalization for Semantic Image Synthesis
http://openaccess.thecvf.com//content/CVPR2022/html/Shi_Retrieval-Based_Spatially_Adaptive_Normalization_for_Semantic_Image_Synthesis_CVPR_2022_paper.html
AUTHORS: Yupeng Shi, Xiao Liu, Yuxiang Wei, Zhongqin Wu, Wangmeng Zuo
HIGHLIGHT: In this paper, we propose a novel normalization module, termed as REtrieval-based Spatially AdaptIve normaLization (RESAIL), for introducing pixel level fine-grained guidance to the normalization architecture.

1047, TITLE: Undoing the Damage of Label Shift for Cross-Domain Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Undoing_the_Damage_of_Label_Shift_for_Cross-Domain_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Yahao Liu, Jinhong Deng, Jiale Tao, Tong Chu, Lixin Duan, Wen Li
HIGHLIGHT: In this paper, we give an in-depth analysis and show that the damage of label shift can be overcome by aligning the data conditional distribution and correcting the posterior probability.

1048, TITLE: GPV-Pose: Category-Level Object Pose Estimation via Geometry-Guided Point-Wise Voting
http://openaccess.thecvf.com//content/CVPR2022/html/Di_GPV-Pose_Category-Level_Object_Pose_Estimation_via_Geometry-Guided_Point-Wise_Voting_CVPR_2022_paper.html
AUTHORS: Yan Di, Ruida Zhang, Zhiqiang Lou, Fabian Manhardt, Xiangyang Ji, Nassir Navab, Federico Tombari
HIGHLIGHT: This is, however, a much more challenging task due to severe intra-class shape variations. To address this issue, we propose GPV-Pose, a novel framework for robust category-level pose estimation, harnessing geometric insights to enhance the learning of category-level pose-sensitive features.

1049, TITLE: Dynamic 3D Gaze From Afar: Deep Gaze Estimation From Temporal Eye-Head-Body Coordination
http://openaccess.thecvf.com//content/CVPR2022/html/Nonaka_Dynamic_3D_Gaze_From_Afar_Deep_Gaze_Estimation_From_Temporal_CVPR_2022_paper.html
AUTHORS: Soma Nonaka, Shohei Nobuhara, Ko Nishino
HIGHLIGHT: We introduce a novel method and dataset for 3D gaze estimation of a freely moving person from a distance, typically in surveillance views.

1050, TITLE: Expressive Talking Head Generation With Granular Audio-Visual Control
http://openaccess.thecvf.com//content/CVPR2022/html/Liang_Expressive_Talking_Head_Generation_With_Granular_Audio-Visual_Control_CVPR_2022_paper.html
AUTHORS: Borong Liang, Yan Pan, Zhizhi Guo, Hang Zhou, Zhibin Hong, Xiaoguang Han, Junyu Han, Jingtuo Liu, Errui Ding, Jingdong Wang
HIGHLIGHT: In this paper, we propose the Granularly Controlled Audio-Visual Talking Heads (GC-AVT), which controls lip movements, head poses, and facial expressions of a talking head in a granular manner.

1051, TITLE: Trustworthy Long-Tailed Classification
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Trustworthy_Long-Tailed_Classification_CVPR_2022_paper.html
AUTHORS: Bolian Li, Zongbo Han, Haining Li, Huazhu Fu, Changqing Zhang
HIGHLIGHT: To address these issues, we propose a Trustworthy Long-tailed Classification (TLC) method to jointly conduct classification and uncertainty estimation to identify hard samples in a multi-expert framework.

1052, TITLE: Primitive3D: 3D Object Dataset Synthesis From Randomly Assembled Primitives
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Primitive3D_3D_Object_Dataset_Synthesis_From_Randomly_Assembled_Primitives_CVPR_2022_paper.html
AUTHORS: Xinke Li, Henghui Ding, Zekun Tong, Yuwei Wu, Yeow Meng Chee
HIGHLIGHT: However, such a dataset is prohibitively expensive in 3D computer vision due to the substantial collection cost. To alleviate this issue, we propose a cost-effective method for automatically generating a large amount of 3D objects with annotations.

1053, TITLE: Mix and Localize: Localizing Sound Sources in Mixtures
http://openaccess.thecvf.com//content/CVPR2022/html/Hu_Mix_and_Localize_Localizing_Sound_Sources_in_Mixtures_CVPR_2022_paper.html
AUTHORS: Xixi Hu, Ziyang Chen, Andrew Owens
HIGHLIGHT: We present a method for simultaneously localizing multiple sound sources within a visual scene.

1054, TITLE: FisherMatch: Semi-Supervised Rotation Regression via Entropy-Based Filtering
http://openaccess.thecvf.com//content/CVPR2022/html/Yin_FisherMatch_Semi-Supervised_Rotation_Regression_via_Entropy-Based_Filtering_CVPR_2022_paper.html
AUTHORS: Yingda Yin, Yingcheng Cai, He Wang, Baoquan Chen
HIGHLIGHT: Inspired by the popular semi-supervised approach, FixMatch, we propose to leverage pseudo label filtering to facilitate the information flow from labeled data to unlabeled data in a teacher-student mutual learning framework.

1055, TITLE: NPBG++: Accelerating Neural Point-Based Graphics

http://openaccess.thecvf.com//content/CVPR2022/html/Rakhimov_NPBG_Accelerating_Neural_Point-Based_Graphics_CVPR_2022_paper.html
AUTHORS: Ruslan Rakhimov, Andrei-Timotei Ardelean, Victor Lempitsky, Evgeny Burnaev
HIGHLIGHT: We present a new system (NPBG++) for the novel view synthesis (NVS) task that achieves high rendering realism with low scene fitting time.

1056, TITLE: SphericGAN: Semi-Supervised Hyper-Spherical Generative Adversarial Networks for Fine-Grained Image Synthesis
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_SphericGAN_Semi-Supervised_Hyper-Spherical_Generative_Adversarial_Networks_for_Fine-Grained_Image_Synthesis_CVPR_2022_paper.html
AUTHORS: Tianyi Chen, Yunfei Zhang, Xiaoyang Huo, Si Wu, Yong Xu, Hau San Wong
HIGHLIGHT: To reduce the dependence of generative models on labeled data, we propose a semi-supervised hyper-spherical GAN for class-conditional fine-grained image generation, and our model is referred to as SphericGAN.

1057, TITLE: HairMapper: Removing Hair From Portraits Using GANs
http://openaccess.thecvf.com//content/CVPR2022/html/Wu_HairMapper_Removing_Hair_From_Portraits_Using_GANs_CVPR_2022_paper.html
AUTHORS: Yiqian Wu, Yong-Liang Yang, Xiaogang Jin
HIGHLIGHT: Removing hair from portrait images is challenging due to the complex occlusions between hair and face, as well as the lack of paired portrait data with/without hair. To this end, we present a dataset and a baseline method for removing hair from portrait images using generative adversarial networks (GANs).

1058, TITLE: Affine Medical Image Registration With Coarse-To-Fine Vision Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Mok_Affine_Medical_Image_Registration_With_Coarse-To-Fine_Vision_Transformer_CVPR_2022_paper.html
AUTHORS: Tony C. W. Mok, Albert C. S. Chung
HIGHLIGHT: In this paper, we present a fast and robust learning-based algorithm, Coarse-to-Fine Vision Transformer (C2FViT), for 3D affine medical image registration.

1059, TITLE: SMPL-A: Modeling Person-Specific Deformable Anatomy
http://openaccess.thecvf.com//content/CVPR2022/html/Guo_SMPL-A_Modeling_Person-Specific_Deformable_Anatomy_CVPR_2022_paper.html
AUTHORS: Hengtao Guo, Benjamin Planche, Meng Zheng, Srikrishna Karanam, Terrence Chen, Ziyan Wu
HIGHLIGHT: To the best of our knowledge, this work is the first to present a learning-based approach to estimate the patient's internal organ deformation for arbitrary human poses in order to assist with radiotherapy and similar medical protocols.

1060, TITLE: Image Dehazing Transformer With Transmission-Aware 3D Position Embedding
http://openaccess.thecvf.com//content/CVPR2022/html/Guo_Image_Dehazing_Transformer_With_Transmission-Aware_3D_Position_Embedding_CVPR_2022_paper.html
AUTHORS: Chun-Le Guo, Qixin Yan, Saeed Anwar, Runmin Cong, Wenqi Ren, Chongyi Li
HIGHLIGHT: The key insight of this study is to investigate how to combine CNN and Transformer for image dehazing.

1061, TITLE: Out-of-Distribution Generalization With Causal Invariant Transformations
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Out-of-Distribution_Generalization_With_Causal_Invariant_Transformations_CVPR_2022_paper.html
AUTHORS: Ruoyu Wang, Mingyang Yi, Zhitang Chen, Shengyu Zhu
HIGHLIGHT: To leverage the generally unknown causal mechanism, existing works assume a linear form of causal feature or require sufficiently many and diverse training domains, which are usually restrictive in practice. In this work, we obviate these assumptions and tackle the OOD problem without explicitly recovering the causal feature.

1062, TITLE: Panoptic-PHNet: Towards Real-Time and High-Precision LiDAR Panoptic Segmentation via Clustering Pseudo Heatmap
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Panoptic-PHNet_Towards_Real-Time_and_High-Precision_LiDAR_Panoptic_Segmentation_via_Clustering_CVPR_2022_paper.html
AUTHORS: Jinke Li, Xiao He, Yang Wen, Yuan Gao, Xiaoqiang Cheng, Dan Zhang
HIGHLIGHT: In this paper, we propose a fast and high-performance LiDAR-based framework, referred to as Panoptic-PHNet, with three attractive aspects: 1) We introduce a clustering pseudo heatmap as a new paradigm, which, followed by a center grouping module, yields instance centers for efficient clustering without object-level learning tasks.

1063, TITLE: Dual-Key Multimodal Backdoors for Visual Question Answering
http://openaccess.thecvf.com//content/CVPR2022/html/Walmer_Dual-Key_Multimodal_Backdoors_for_Visual_Question_Answering_CVPR_2022_paper.html
AUTHORS: Matthew Walmer, Karan Sikka, Indranil Sur, Abhinav Shrivastava, Susmit Jha
HIGHLIGHT: In this work, we show that multimodal networks are vulnerable to a novel type of attack that we refer to as Dual-Key Multimodal Backdoors.

1064, TITLE: A Differentiable Two-Stage Alignment Scheme for Burst Image Reconstruction With Large Shift
http://openaccess.thecvf.com//content/CVPR2022/html/Guo_A_Differentiable_Two-Stage_Alignment_Scheme_for_Burst_Image_Reconstruction_With_CVPR_2022_paper.html
AUTHORS: Shi Guo, Xi Yang, Jianqi Ma, Gaofeng Ren, Lei Zhang

HIGHLIGHT:	In addition, the higher resolution (e.g., 4K) of modern imaging devices results in larger displacement between frames. To address these challenges, we design a differentiable two-stage alignment scheme sequentially in patch and pixel level for effective JDD-B.

1065, TITLE:	Unifying Panoptic Segmentation for Autonomous Driving
http://openaccess.thecvf.com//content/CVPR2022/html/Zendel_Unifying_Panoptic_Segmentation_for_Autonomous_Driving_CVPR_2022_paper.html
AUTHORS:	Oliver Zendel, Matthias Schörghuber, Bernhard Rainer, Markus Murschitz, Csaba Beleznai
HIGHLIGHT:	This paper aims to improve panoptic segmentation for real-world applications in three ways.

1066, TITLE:	Learning Motion-Dependent Appearance for High-Fidelity Rendering of Dynamic Humans From a Single Camera
http://openaccess.thecvf.com//content/CVPR2022/html/Yoon_Learning_Motion-Dependent_Appearance_for_High-Fidelity_Rendering_of_Dynamic_Humans_From_CVPR_2022_paper.html
AUTHORS:	Jae Shin Yoon, Duygu Ceylan, Tuanfeng Y. Wang, Jingwan Lu, Jimei Yang, Zhixin Shu, Hyun Soo Park
HIGHLIGHT:	In this paper, we present a compact motion representation by enforcing equivariance---a representation is expected to be transformed in the way that the pose is transformed.

1067, TITLE:	On the Road to Online Adaptation for Semantic Image Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Volpi_On_the_Road_to_Online_Adaptation_for_Semantic_Image_Segmentation_CVPR_2022_paper.html
AUTHORS:	Riccardo Volpi, Pau De Jorge, Diane Larlus, Gabriela Csurka
HIGHLIGHT:	We propose a new problem formulation and a corresponding evaluation framework to advance research on unsupervised domain adaptation for semantic image segmentation.

1068, TITLE:	Deformable ProtoPNet: An Interpretable Image Classifier Using Deformable Prototypes
http://openaccess.thecvf.com//content/CVPR2022/html/Donnelly_Deformable_ProtoPNet_An_Interpretable_Image_Classifier_Using_Deformable_Prototypes_CVPR_2022_paper.html
AUTHORS:	Jon Donnelly, Alina Jade Barnett, Chaofan Chen
HIGHLIGHT:	We present a deformable prototypical part network (Deformable ProtoPNet), an interpretable image classifier that integrates the power of deep learning and the interpretability of case-based reasoning.

1069, TITLE:	Context-Aware Sequence Alignment Using 4D Skeletal Augmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Kwon_Context-Aware_Sequence_Alignment_Using_4D_Skeletal_Augmentation_CVPR_2022_paper.html
AUTHORS:	Taein Kwon, Bugra Tekin, Siyu Tang, Marc Pollefeys
HIGHLIGHT:	In this work, based on off-the-shelf human pose estimators, we propose a novel context-aware self-supervised learning architecture to align sequences of actions.

1070, TITLE:	Perturbed and Strict Mean Teachers for Semi-Supervised Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Perturbed_and_Strict_Mean_Teachers_for_Semi-Supervised_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS:	Yuyuan Liu, Yu Tian, Yuanhong Chen, Fengbei Liu, Vasileios Belagiannis, Gustavo Carneiro
HIGHLIGHT:	In this paper, we address the prediction accuracy problem of consistency learning methods with novel extensions of the mean-teacher (MT) model, which include a new auxiliary teacher, and the replacement of MT's mean square error (MSE) by a stricter confidence-weighted cross-entropy (Conf-CE) loss.

1071, TITLE:	Motion-Modulated Temporal Fragment Alignment Network for Few-Shot Action Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Wu_Motion-Modulated_Temporal_Fragment_Alignment_Network_for_Few-Shot_Action_Recognition_CVPR_2022_paper.html
AUTHORS:	Jiamin Wu, Tianzhu Zhang, Zhe Zhang, Feng Wu, Yongdong Zhang
HIGHLIGHT:	While the majority of FSL models focus on image classification, the extension to action recognition is rather challenging due to the additional temporal dimension in videos. To address this issue, we propose an end-to-end Motion-modulated Temporal Fragment Alignment Network (MTFAN) by jointly exploring the task-specific motion modulation and the multi-level temporal fragment alignment for Few-Shot Action Recognition (FSAR).

1072, TITLE:	Focal Sparse Convolutional Networks for 3D Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Focal_Sparse_Convolutional_Networks_for_3D_Object_Detection_CVPR_2022_paper.html
AUTHORS:	Yukang Chen, Yanwei Li, Xiangyu Zhang, Jian Sun, Jiaya Jia
HIGHLIGHT:	In this paper, we introduce two new modules to enhance the capability of Sparse CNNs, both are based on making feature sparsity learnable with position-wise importance prediction.

1073, TITLE:	Masked Autoencoders Are Scalable Vision Learners
http://openaccess.thecvf.com//content/CVPR2022/html/He_Masked_Autoencoders_Are_Scalable_Vision_Learners_CVPR_2022_paper.html
AUTHORS:	Kaiming He, Xinlei Chen, Saining Xie, Yanghao Li, Piotr Dollár, Ross Girshick
HIGHLIGHT:	This paper shows that masked autoencoders (MAE) are scalable self-supervised learners for computer vision.

1074, TITLE:	Point-BERT: Pre-Training 3D Point Cloud Transformers With Masked Point Modeling

http://openaccess.thecvf.com//content/CVPR2022/html/Yu_Point-BERT_Pre-Training_3D_Point_Cloud_Transformers_With_Masked_Point_Modeling_CVPR_2022_paper.html
AUTHORS: Xumin Yu, Lulu Tang, Yongming Rao, Tiejun Huang, Jie Zhou, Jiwen Lu
HIGHLIGHT: We present Point-BERT, a novel paradigm for learning Transformers to generalize the concept of BERT onto 3D point cloud.

1075, TITLE: Nested Collaborative Learning for Long-Tailed Visual Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Nested_Collaborative_Learning_for_Long-Tailed_Visual_Recognition_CVPR_2022_paper.html
AUTHORS: Jun Li, Zichang Tan, Jun Wan, Zhen Lei, Guodong Guo
HIGHLIGHT: The networks trained on the long-tailed dataset vary remarkably, despite the same training settings, which shows the great uncertainty in long-tailed learning. To alleviate the uncertainty, we propose a Nested Collaborative Learning (NCL), which tackles the problem by collaboratively learning multiple experts together.

1076, TITLE: Crowd Counting in the Frequency Domain
http://openaccess.thecvf.com//content/CVPR2022/html/Shu_Crowd_Counting_in_the_Frequency_Domain_CVPR_2022_paper.html
AUTHORS: Weibo Shu, Jia Wan, Kay Chen Tan, Sam Kwong, Antoni B. Chan
HIGHLIGHT: By transforming the density map into the frequency domain and using the nice properties of the characteristic function, we propose a novel method that is simple, effective, and efficient.

1077, TITLE: Restormer: Efficient Transformer for High-Resolution Image Restoration
http://openaccess.thecvf.com//content/CVPR2022/html/Zamir_Restormer_Efficient_Transformer_for_High-Resolution_Image_Restoration_CVPR_2022_paper.html
AUTHORS: Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang
HIGHLIGHT: In this work, we propose an efficient Transformer model by making several key designs in the building blocks (multi-head attention and feed-forward network) such that it can capture long-range pixel interactions, while still remaining applicable to large images.

1078, TITLE: STRPM: A Spatiotemporal Residual Predictive Model for High-Resolution Video Prediction
http://openaccess.thecvf.com//content/CVPR2022/html/Chang_STRPM_A_Spatiotemporal_Residual_Predictive_Model_for_High-Resolution_Video_Prediction_CVPR_2022_paper.html
AUTHORS: Zheng Chang, Xinfeng Zhang, Shanshe Wang, Siwei Ma, Wen Gao
HIGHLIGHT: In this paper, we propose a Spatiotemporal Residual Predictive Model (STRPM) for high-resolution video prediction.

1079, TITLE: Learning From Untrimmed Videos: Self-Supervised Video Representation Learning With Hierarchical Consistency
http://openaccess.thecvf.com//content/CVPR2022/html/Qing_Learning_From_Untrimmed_Videos_Self-Supervised_Video_Representation_Learning_With_Hierarchical_CVPR_2022_paper.html
AUTHORS: Zhiwu Qing, Shiwei Zhang, Ziyuan Huang, Yi Xu, Xiang Wang, Mingqian Tang, Changxin Gao, Rong Jin, Nong Sang
HIGHLIGHT: In this work, we aim to learn representations by leveraging more abundant information in untrimmed videos.

1080, TITLE: Aladdin: Joint Atlas Building and Diffeomorphic Registration Learning With Pairwise Alignment
http://openaccess.thecvf.com//content/CVPR2022/html/Ding_Aladdin_Joint_Atlas_Building_and_Diffeomorphic_Registration_Learning_With_Pairwise_CVPR_2022_paper.html
AUTHORS: Zhipeng Ding, Marc Niethammer
HIGHLIGHT: This work explores using a convolutional neural network (CNN) to jointly predict the atlas and a stationary velocity field (SVF) parameterization for diffeomorphic image registration with respect to the atlas.

1081, TITLE: IFRNet: Intermediate Feature Refine Network for Efficient Frame Interpolation
http://openaccess.thecvf.com//content/CVPR2022/html/Kong_IFRNet_Intermediate_Feature_Refine_Network_for_Efficient_Frame_Interpolation_CVPR_2022_paper.html
AUTHORS: Lingtong Kong, Boyuan Jiang, Donghao Luo, Wenqing Chu, Xiaoming Huang, Ying Tai, Chengjie Wang, Jie Yang
HIGHLIGHT: In this work, we devise an efficient encoder-decoder based network, termed IFRNet, for fast intermediate frame synthesizing.

1082, TITLE: Large Loss Matters in Weakly Supervised Multi-Label Classification
http://openaccess.thecvf.com//content/CVPR2022/html/Kim_Large_Loss_Matters_in_Weakly_Supervised_Multi-Label_Classification_CVPR_2022_paper.html
AUTHORS: Youngwook Kim, Jae Myung Kim, Zeynep Akata, Jungwoo Lee
HIGHLIGHT: That is, the model first learns the representation of clean labels, and then starts memorizing noisy labels. Based on this finding, we propose novel methods for WSML which reject or correct the large loss samples to prevent model from memorizing the noisy label.

1083, TITLE: Toward Practical Monocular Indoor Depth Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Wu_Toward_Practical_Monocular_Indoor_Depth_Estimation_CVPR_2022_paper.html
AUTHORS: Cho-Ying Wu, Jialiang Wang, Michael Hall, Ulrich Neumann, Shuochen Su

HIGHLIGHT:	To obtain more robustness, we propose a structure distillation approach to learn knacks from an off-the-shelf relative depth estimator that produces structured but metric-agnostic depth.

1084, TITLE:	Attention Concatenation Volume for Accurate and Efficient Stereo Matching
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_Attention_Concatenation_Volume_for_Accurate_and_Efficient_Stereo_Matching_CVPR_2022_paper.html
AUTHORS:	Gangwei Xu, Junda Cheng, Peng Guo, Xin Yang
HIGHLIGHT:	In this paper, we present a novel cost volume construction method which generates attention weights from correlation clues to suppress redundant information and enhance matching-related information in the concatenation volume.

1085, TITLE:	Learning Distinctive Margin Toward Active Domain Adaptation
http://openaccess.thecvf.com//content/CVPR2022/html/Xie_Learning_Distinctive_Margin_Toward_Active_Domain_Adaptation_CVPR_2022_paper.html
AUTHORS:	Ming Xie, Yuxi Li, Yabiao Wang, Zekun Luo, Zhenye Gan, Zhongyi Sun, Mingmin Chi, Chengjie Wang, Pei Wang
HIGHLIGHT:	In this work, we propose a concise but effective ADA method called Select-by-Distinctive-Margin (SDM), which consists of a maximum margin loss and a margin sampling algorithm for data selection.

1086, TITLE:	Zero-Query Transfer Attacks on Context-Aware Object Detectors
http://openaccess.thecvf.com//content/CVPR2022/html/Cai_Zero-Query_Transfer_Attacks_on_Context-Aware_Object_Detectors_CVPR_2022_paper.html
AUTHORS:	Zikui Cai, Shantanu Rane, Alejandro E. Brito, Chengyu Song, Srikanth V. Krishnamurthy, Amit K. Roy-Chowdhury, M. Salman Asif
HIGHLIGHT:	We present the first approach for generating context-consistent adversarial attacks that can evade the context-consistency check of black-box object detectors operating on complex, natural scenes.

1087, TITLE:	Neural Inertial Localization
http://openaccess.thecvf.com//content/CVPR2022/html/Herath_Neural_Inertial_Localization_CVPR_2022_paper.html
AUTHORS:	Sachini Herath, David Caruso, Chen Liu, Yufan Chen, Yasutaka Furukawa
HIGHLIGHT:	This paper proposes the inertial localization problem, the task of estimating the absolute location from a sequence of inertial sensor measurements.

1088, TITLE:	Speed Up Object Detection on Gigapixel-Level Images With Patch Arrangement
http://openaccess.thecvf.com//content/CVPR2022/html/Fan_Speed_Up_Object_Detection_on_Gigapixel-Level_Images_With_Patch_Arrangement_CVPR_2022_paper.html
AUTHORS:	Jiahao Fan, Huabin Liu, Wenjie Yang, John See, Aixin Zhang, Weiyao Lin
HIGHLIGHT:	Different from them, we devise a novel patch arrangement framework for fast object detection on gigapixel-level images.

1089, TITLE:	Finding Fallen Objects via Asynchronous Audio-Visual Integration
http://openaccess.thecvf.com//content/CVPR2022/html/Gan_Finding_Fallen_Objects_via_Asynchronous_Audio-Visual_Integration_CVPR_2022_paper.html
AUTHORS:	Chuang Gan, Yi Gu, Siyuan Zhou, Jeremy Schwartz, Seth Alter, James Traer, Dan Gutfreund, Joshua B. Tenenbaum, Josh H. McDermott, Antonio Torralba
HIGHLIGHT:	In this paper, we introduce a setting in which to study multi-modal object localization in 3D virtual environments.

1090, TITLE:	Learning sRGB-to-Raw-RGB De-Rendering With Content-Aware Metadata
http://openaccess.thecvf.com//content/CVPR2022/html/Nam_Learning_sRGB-to-Raw-RGB_De-Rendering_With_Content-Aware_Metadata_CVPR_2022_paper.html
AUTHORS:	Seonghyeon Nam, Abhijith Punnappurath, Marcus A. Brubaker, Michael S. Brown
HIGHLIGHT:	This paper shows how to improve the de-rendering results by jointly learning sampling and reconstruction.

1091, TITLE:	GraftNet: Towards Domain Generalized Stereo Matching With a Broad-Spectrum and Task-Oriented Feature
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_GraftNet_Towards_Domain_Generalized_Stereo_Matching_With_a_Broad-Spectrum_and_CVPR_2022_paper.html
AUTHORS:	Biyang Liu, Huimin Yu, Guodong Qi
HIGHLIGHT:	In this paper, we propose to leverage the feature of a model trained on large-scale datasets to deal with the domain shift since it has seen various styles of images.

1092, TITLE:	Towards Total Recall in Industrial Anomaly Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Roth_Towards_Total_Recall_in_Industrial_Anomaly_Detection_CVPR_2022_paper.html
AUTHORS:	Karsten Roth, Latha Pemula, Joaquin Zepeda, Bernhard Schölkopf, Thomas Brox, Peter Gehler
HIGHLIGHT:	The best peforming approaches combine embeddings from ImageNet models with an outlier detection model. In this paper, we extend on this line of work and propose PatchCore, which uses a maximally representative memory bank of nominal patch-features.

1093, TITLE:	DTA: Physical Camouflage Attacks Using Differentiable Transformation Network

http://openaccess.thecvf.com//content/CVPR2022/html/Suryanto_DTA_Physical_Camouflage_Attacks_Using_Differentiable_Transformation_Network_CVPR_2022_paper.html
AUTHORS: Naufal Suryanto, Yongsu Kim, Hyoeun Kang, Harashta Tatimma Larasati, Youngyeo Yun, Thi-Thu-Huong Le, Hunmin Yang, Se-Yoon Oh, Howon Kim
HIGHLIGHT: In this paper, we propose the Differentiable Transformation Attack (DTA), a framework for generating a robust physical adversarial pattern on a target object to camouflage it against object detection models with a wide range of transformations.

1094, TITLE: Neural Recognition of Dashed Curves With Gestalt Law of Continuity
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Neural_Recognition_of_Dashed_Curves_With_Gestalt_Law_of_Continuity_CVPR_2022_paper.html
AUTHORS: Hanyuan Liu, Chengze Li, Xueting Liu, Tien-Tsin Wong
HIGHLIGHT: In this paper, we propose an innovative Transformer-based framework to recognize dashed curves based on both high-level features and low-level clues.

1095, TITLE: Semi-Supervised Object Detection via Multi-Instance Alignment With Global Class Prototypes
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Semi-Supervised_Object_Detection_via_Multi-Instance_Alignment_With_Global_Class_Prototypes_CVPR_2022_paper.html
AUTHORS: Aoxue Li, Peng Yuan, Zhenguo Li
HIGHLIGHT: To make full use of labeled data, we propose a Multi-instance Alignment model which enhances the prediction consistency based on Global Class Prototypes (MA-GCP).

1096, TITLE: HODOR: High-Level Object Descriptors for Object Re-Segmentation in Video Learned From Static Images
http://openaccess.thecvf.com//content/CVPR2022/html/Athar_HODOR_High-Level_Object_Descriptors_for_Object_Re-Segmentation_in_Video_Learned_CVPR_2022_paper.html
AUTHORS: Ali Athar, Jonathon Luiten, Alexander Hermans, Deva Ramanan, Bastian Leibe
HIGHLIGHT: This requires a large amount of densely annotated video data, which is costly to annotate, and largely redundant since frames within a video are highly correlated. In light of this, we propose HODOR: a novel method that tackles VOS by effectively leveraging annotated static images for understanding object appearance and scene context.

1097, TITLE: Point Cloud Color Constancy
http://openaccess.thecvf.com//content/CVPR2022/html/Xing_Point_Cloud_Color_Constancy_CVPR_2022_paper.html
AUTHORS: Xiaoyan Xing, Yanlin Qian, Sibo Feng, Yuhan Dong, Jiří Matas
HIGHLIGHT: In this paper, we present Point Cloud Color Constancy, in short PCCC, an illumination chromaticity estimation algorithm exploiting a point cloud.

1098, TITLE: VGSE: Visually-Grounded Semantic Embeddings for Zero-Shot Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_VGSE_Visually-Grounded_Semantic_Embeddings_for_Zero-Shot_Learning_CVPR_2022_paper.html
AUTHORS: Wenjia Xu, Yongqin Xian, Jiuniu Wang, Bernt Schiele, Zeynep Akata
HIGHLIGHT: We propose to discover semantic embeddings containing discriminative visual properties for zero-shot learning, without requiring any human annotation.

1099, TITLE: Catching Both Gray and Black Swans: Open-Set Supervised Anomaly Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Ding_Catching_Both_Gray_and_Black_Swans_Open-Set_Supervised_Anomaly_Detection_CVPR_2022_paper.html
AUTHORS: Choubo Ding, Guansong Pang, Chunhua Shen
HIGHLIGHT: We propose a novel approach that learns disentangled representations of abnormalities illustrated by seen anomalies, pseudo anomalies, and latent residual anomalies (i.e., samples that have unusual residuals compared to the normal data in a latent space), with the last two abnormalities designed to detect unseen anomalies.

1100, TITLE: MLSLT: Towards Multilingual Sign Language Translation
http://openaccess.thecvf.com//content/CVPR2022/html/Yin_MLSLT_Towards_Multilingual_Sign_Language_Translation_CVPR_2022_paper.html
AUTHORS: Aoxiong Yin, Zhou Zhao, Weike Jin, Meng Zhang, Xingshan Zeng, Xiaofei He
HIGHLIGHT: However, such models are inefficient in building multilingual sign language translation systems. To solve this problem, we introduce the multilingual sign language translation (MSLT) task.

1101, TITLE: Towards an End-to-End Framework for Flow-Guided Video Inpainting
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Towards_an_End-to-End_Framework_for_Flow-Guided_Video_Inpainting_CVPR_2022_paper.html
AUTHORS: Zhen Li, Cheng-Ze Lu, Jianhua Qin, Chun-Le Guo, Ming-Ming Cheng
HIGHLIGHT: In this paper, we propose an End-to-End framework for Flow-Guided Video Inpainting through elaborately designed three trainable modules, namely, flow completion, feature propagation, and content hallucination modules.

1102, TITLE: Contrastive Test-Time Adaptation
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Contrastive_Test-Time_Adaptation_CVPR_2022_paper.html
AUTHORS: Dian Chen, Dequan Wang, Trevor Darrell, Sayna Ebrahimi
HIGHLIGHT: We propose a novel way to leverage self-supervised contrastive learning to facilitate target feature learning, along with an online pseudo labeling scheme with refinement that significantly denoises pseudo labels.

1103, TITLE:	Multimodal Colored Point Cloud to Image Alignment
http://openaccess.thecvf.com//content/CVPR2022/html/Rotstein_Multimodal_Colored_Point_Cloud_to_Image_Alignment_CVPR_2022_paper.html
AUTHORS:	Noam Rotstein, Amit Bracha, Ron Kimmel
HIGHLIGHT:	A major challenge in acquiring such ground truth data is the accurate alignment between RGB images and the point cloud measured by a depth scanner. To overcome this difficulty, we consider a differential optimization method that aligns a colored point cloud with a given color image through iterative geometric and color matching.

1104, TITLE:	MotionAug: Augmentation With Physical Correction for Human Motion Prediction
http://openaccess.thecvf.com//content/CVPR2022/html/Maeda_MotionAug_Augmentation_With_Physical_Correction_for_Human_Motion_Prediction_CVPR_2022_paper.html
AUTHORS:	Takahiro Maeda, Norimichi Ukita
HIGHLIGHT:	This paper presents a motion data augmentation scheme incorporating motion synthesis encouraging diversity and motion correction imposing physical plausibility.

1105, TITLE:	Active Teacher for Semi-Supervised Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Mi_Active_Teacher_for_Semi-Supervised_Object_Detection_CVPR_2022_paper.html
AUTHORS:	Peng Mi, Jianghang Lin, Yiyi Zhou, Yunhang Shen, Gen Luo, Xiaoshuai Sun, Liujuan Cao, Rongrong Fu, Qiang Xu, Rongrong Ji
HIGHLIGHT:	In this paper, we study teacher-student learning from the perspective of data initialization and propose a novel algorithm called Active Teacher for semi-supervised object detection (SSOD).

1106, TITLE:	CrossLoc: Scalable Aerial Localization Assisted by Multimodal Synthetic Data
http://openaccess.thecvf.com//content/CVPR2022/html/Yan_CrossLoc_Scalable_Aerial_Localization_Assisted_by_Multimodal_Synthetic_Data_CVPR_2022_paper.html
AUTHORS:	Qi Yan, Jianhao Zheng, Simon Reding, Shanci Li, Iordan Doytchinov
HIGHLIGHT:	We present a visual localization system that learns to estimate camera poses in the real world with the help of synthetic data.

1107, TITLE:	Audio-Adaptive Activity Recognition Across Video Domains
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Audio-Adaptive_Activity_Recognition_Across_Video_Domains_CVPR_2022_paper.html
AUTHORS:	Yunhua Zhang, Hazel Doughty, Ling Shao, Cees G. M. Snoek
HIGHLIGHT:	We propose an audio-adaptive encoder and associated learning methods that discriminatively adjust the visual feature representation as well as addressing shifts in the semantic distribution.

1108, TITLE:	Collaborative Learning for Hand and Object Reconstruction With Attention-Guided Graph Convolution
http://openaccess.thecvf.com//content/CVPR2022/html/Tse_Collaborative_Learning_for_Hand_and_Object_Reconstruction_With_Attention-Guided_Graph_CVPR_2022_paper.html
AUTHORS:	Tze Ho Elden Tse, Kwang In Kim, AlesÌ† Leonardis, Hyung Jin Chang
HIGHLIGHT:	Specifically, we transfer hand mesh information to the object branch and vice versa for the hand branch.

1109, TITLE:	On Learning Contrastive Representations for Learning With Noisy Labels
http://openaccess.thecvf.com//content/CVPR2022/html/Yi_On_Learning_Contrastive_Representations_for_Learning_With_Noisy_Labels_CVPR_2022_paper.html
AUTHORS:	Li Yi, Sheng Liu, Qi She, A. Ian McLeod, Boyu Wang
HIGHLIGHT:	To address this issue, we focus on learning robust contrastive representations of data on which the classifier is hard to memorize the label noise under the CE loss. We propose a novel contrastive regularization function to learn such representations over noisy data where the label noise does not dominate the representation learning.

1110, TITLE:	Unsupervised Deraining: Where Contrastive Learning Meets Self-Similarity
http://openaccess.thecvf.com//content/CVPR2022/html/Ye_Unsupervised_Deraining_Where_Contrastive_Learning_Meets_Self-Similarity_CVPR_2022_paper.html
AUTHORS:	Yuntong Ye, Changfeng Yu, Yi Chang, Lin Zhu, Xi-Le Zhao, Luxin Yan, Yonghong Tian
HIGHLIGHT:	In this work, we propose a novel non-local contrastive learning (NLCL) method for unsupervised image deraining.

1111, TITLE:	Modeling Indirect Illumination for Inverse Rendering
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Modeling_Indirect_Illumination_for_Inverse_Rendering_CVPR_2022_paper.html
AUTHORS:	Yuanqing Zhang, Jiaming Sun, Xingyi He, Huan Fu, Rongfei Jia, Xiaowei Zhou
HIGHLIGHT:	In this paper, we propose a novel approach to efficiently recovering spatially-varying indirect illumination.

1112, TITLE:	BACON: Band-Limited Coordinate Networks for Multiscale Scene Representation
http://openaccess.thecvf.com//content/CVPR2022/html/Lindell_BACON_Band-Limited_Coordinate_Networks_for_Multiscale_Scene_Representation_CVPR_2022_paper.html
AUTHORS:	David B. Lindell, Dave Van Veen, Jeong Joon Park, Gordon Wetzstein
HIGHLIGHT:	We introduce band-limited coordinate networks (BACON), a network architecture with an analytical Fourier spectrum.

1113, TITLE: Regional Semantic Contrast and Aggregation for Weakly Supervised Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Zhou_Regional_Semantic_Contrast_and_Aggregation_for_Weakly_Supervised_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Tianfei Zhou, Meijie Zhang, Fang Zhao, Jianwu Li
HIGHLIGHT: In particular, we propose regional semantic contrast and aggregation (RCA).

1114, TITLE: Class Re-Activation Maps for Weakly-Supervised Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Class_Re-Activation_Maps_for_Weakly-Supervised_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Zhaozheng Chen, Tan Wang, Xiongwei Wu, Xian-Sheng Hua, Hanwang Zhang, Qianru Sun
HIGHLIGHT: As a result, given a class, its hot CAM pixels may wrongly invade the area belonging to other classes, or the non-hot ones may be actually a part of the class. To this end, we introduce an embarrassingly simple yet surprisingly effective method: Reactivating the converged CAM with BCE by using softmax cross-entropy loss (SCE), dubbed ReCAM.

1115, TITLE: TransWeather: Transformer-Based Restoration of Images Degraded by Adverse Weather Conditions
http://openaccess.thecvf.com//content/CVPR2022/html/Valanarasu_TransWeather_Transformer-Based_Restoration_of_Images_Degraded_by_Adverse_Weather_Conditions_CVPR_2022_paper.html
AUTHORS: Jeya Maria Jose Valanarasu, Rajeev Yasarla, Vishal M. Patel
HIGHLIGHT: In this work, we focus on developing an efficient solution for the all adverse weather removal problem.

1116, TITLE: Merry Go Round: Rotate a Frame and Fool a DNN
http://openaccess.thecvf.com//content/CVPR2022/html/Thapar_Merry_Go_Round_Rotate_a_Frame_and_Fool_a_DNN_CVPR_2022_paper.html
AUTHORS: Daksh Thapar, Aditya Nigam, Chetan Arora
HIGHLIGHT: In this paper we make a key novel suggestion to use perturbation in optical flow to carry out AAs on a video analysis system.

1117, TITLE: H2FA R-CNN: Holistic and Hierarchical Feature Alignment for Cross-Domain Weakly Supervised Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_H2FA_R-CNN_Holistic_and_Hierarchical_Feature_Alignment_for_Cross-Domain_Weakly_CVPR_2022_paper.html
AUTHORS: Yunqiu Xu, Yifan Sun, Zongxin Yang, Jiaxu Miao, Yi Yang
HIGHLIGHT: Existing methods usually focus on partial detection components for domain alignment. In contrast, this paper considers that all the detection components are important and proposes a Holistic and Hierarchical Feature Alignment (H^2FA) R-CNN.

1118, TITLE: Modeling sRGB Camera Noise With Normalizing Flows
http://openaccess.thecvf.com//content/CVPR2022/html/Kousha_Modeling_sRGB_Camera_Noise_With_Normalizing_Flows_CVPR_2022_paper.html
AUTHORS: Shayan Kousha, Ali Maleky, Michael S. Brown, Marcus A. Brubaker
HIGHLIGHT: We propose a new sRGB-domain noise model based on normalizing flows that is capable of learning the complex noise distribution found in sRGB images under various ISO levels.

1119, TITLE: A ConvNet for the 2020s
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_A_ConvNet_for_the_2020s_CVPR_2022_paper.html
AUTHORS: Zhuang Liu, Hanzi Mao, Chao-Yuan Wu, Christoph Feichtenhofer, Trevor Darrell, Saining Xie
HIGHLIGHT: In this work, we reexamine the design spaces and test the limits of what a pure ConvNet can achieve.

1120, TITLE: Reference-Based Video Super-Resolution Using Multi-Camera Video Triplets
http://openaccess.thecvf.com//content/CVPR2022/html/Lee_Reference-Based_Video_Super-Resolution_Using_Multi-Camera_Video_Triplets_CVPR_2022_paper.html
AUTHORS: Junyong Lee, Myeonghee Lee, Sunghyun Cho, Seungyong Lee
HIGHLIGHT: We propose the first reference-based video super-resolution (RefVSR) approach that utilizes reference videos for high-fidelity results.

1121, TITLE: Self-Supervised Image Representation Learning With Geometric Set Consistency
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Self-Supervised_Image_Representation_Learning_With_Geometric_Set_Consistency_CVPR_2022_paper.html
AUTHORS: Nenglun Chen, Lei Chu, Hao Pan, Yan Lu, Wenping Wang
HIGHLIGHT: We propose a method for self-supervised image representation learning under the guidance of 3D geometric consistency.

1122, TITLE: Deep Anomaly Discovery From Unlabeled Videos via Normality Advantage and Self-Paced Refinement
http://openaccess.thecvf.com//content/CVPR2022/html/Yu_Deep_Anomaly_Discovery_From_Unlabeled_Videos_via_Normality_Advantage_and_CVPR_2022_paper.html
AUTHORS: Guang Yu, Siqi Wang, Zhiping Cai, Xinwang Liu, Chuanfu Xu, Chengkun Wu
HIGHLIGHT: This paper proposes a full deep neural network (DNN) based solution that can realize highly effective UVAD.

1123, TITLE: P3Depth: Monocular Depth Estimation With a Piecewise Planarity Prior

119

http://openaccess.thecvf.com//content/CVPR2022/html/Patil_P3Depth_Monocular_Depth_Estimation_With_a_Piecewise_Planarity_Prior_CVPR_2022_paper.html
AUTHORS: Vaishakh Patil, Christos Sakaridis, Alexander Liniger, Luc Van Gool
HIGHLIGHT: Based on knowledge about the high regularity of real 3D scenes, we propose a method that learns to selectively leverage information from coplanar pixels to improve the predicted depth.

1124, TITLE: GEN-VLKT: Simplify Association and Enhance Interaction Understanding for HOI Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Liao_GEN-VLKT_Simplify_Association_and_Enhance_Interaction_Understanding_for_HOI_Detection_CVPR_2022_paper.html
AUTHORS: Yue Liao, Aixi Zhang, Miao Lu, Yongliang Wang, Xiaobo Li, Si Liu
HIGHLIGHT: In this paper, we reveal and address the disadvantages of the conventional query-driven HOI detectors from the two aspects.

1125, TITLE: Simple Multi-Dataset Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Zhou_Simple_Multi-Dataset_Detection_CVPR_2022_paper.html
AUTHORS: Xingyi Zhou, Vladlen Koltun, Philipp Krähenbühl
HIGHLIGHT: In this paper, we present a simple method for training a unified detector on multiple large-scale datasets.

1126, TITLE: MLP-3D: A MLP-Like 3D Architecture With Grouped Time Mixing
http://openaccess.thecvf.com//content/CVPR2022/html/Qiu_MLP-3D_A_MLP-Like_3D_Architecture_With_Grouped_Time_Mixing_CVPR_2022_paper.html
AUTHORS: Zhaofan Qiu, Ting Yao, Chong-Wah Ngo, Tao Mei
HIGHLIGHT: In this paper, we present MLP-3D networks, a novel MLP-like 3D architecture for video recognition.

1127, TITLE: Proactive Image Manipulation Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Asnani_Proactive_Image_Manipulation_Detection_CVPR_2022_paper.html
AUTHORS: Vishal Asnani, Xi Yin, Tal Hassner, Sijia Liu, Xiaoming Liu
HIGHLIGHT: By contrast, we propose a proactive scheme to image manipulation detection.

1128, TITLE: Sketch3T: Test-Time Training for Zero-Shot SBIR
http://openaccess.thecvf.com//content/CVPR2022/html/Sain_Sketch3T_Test-Time_Training_for_Zero-Shot_SBIR_CVPR_2022_paper.html
AUTHORS: Aneeshan Sain, Ayan Kumar Bhunia, Vaishnav Potlapalli, Pinaki Nath Chowdhury, Tao Xiang, Yi-Zhe Song
HIGHLIGHT: Zero-shot sketch-based image retrieval typically asks for a trained model to be applied as is to unseen categories. In this paper, we question to argue that this setup by definition is not compatible with the inherent abstract and subjective nature of sketches -- the model might transfer well to new categories, but will not understand sketches existing in different test-time distribution as a result.

1129, TITLE: BANMo: Building Animatable 3D Neural Models From Many Casual Videos
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_BANMo_Building_Animatable_3D_Neural_Models_From_Many_Casual_Videos_CVPR_2022_paper.html
AUTHORS: Gengshan Yang, Minh Vo, Natalia Neverova, Deva Ramanan, Andrea Vedaldi, Hanbyul Joo
HIGHLIGHT: Our key insight is to merge three schools of thought: (1) classic deformable shape models that make use of articulated bones and blend skinning, (2) canonical embeddings that establish correspondences between pixels and a canonical 3D model, and (3) volumetric neural radiance fields (NeRFs) that are amenable to gradient-based optimization.

1130, TITLE: StyTr2: Image Style Transfer With Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Deng_StyTr2_Image_Style_Transfer_With_Transformers_CVPR_2022_paper.html
AUTHORS: Yingying Deng, Fan Tang, Weiming Dong, Chongyang Ma, Xingjia Pan, Lei Wang, Changsheng Xu
HIGHLIGHT: Therefore, traditional neural style transfer methods face biased content representation. To address this critical issue, we take long-range dependencies of input images into account for image style transfer by proposing a transformer-based approach called StyTr^2.

1131, TITLE: Towards Discriminative Representation: Multi-View Trajectory Contrastive Learning for Online Multi-Object Tracking
http://openaccess.thecvf.com//content/CVPR2022/html/Yu_Towards_Discriminative_Representation_Multi-View_Trajectory_Contrastive_Learning_for_Online_Multi-Object_CVPR_2022_paper.html
AUTHORS: En Yu, Zhuoling Li, Shoudong Han
HIGHLIGHT: Although this strategy is effective, it fails to fully exploit the information contained in a whole trajectory. To this end, we propose a strategy, namely multi-view trajectory contrastive learning, in which each trajectory is represented as a center vector.

1132, TITLE: Global Matching With Overlapping Attention for Optical Flow Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Zhao_Global_Matching_With_Overlapping_Attention_for_Optical_Flow_Estimation_CVPR_2022_paper.html
AUTHORS: Shiyu Zhao, Long Zhao, Zhixing Zhang, Enyu Zhou, Dimitris Metaxas
HIGHLIGHT: In this paper, inspired by the traditional matching-optimization methods where matching is introduced to handle large displacements before energy-based optimizations, we introduce a simple but effective global matching step before the direct regression and develop a learning-based matching-optimization framework, namely GMFlowNet.

1133, TITLE: Language As Queries for Referring Video Object Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Wu_Language_As_Queries_for_Referring_Video_Object_Segmentation_CVPR_2022_paper.html
AUTHORS: Jiannan Wu, Yi Jiang, Peize Sun, Zehuan Yuan, Ping Luo
HIGHLIGHT: In this work, we propose a simple and unified framework built upon Transformer, termed ReferFormer.

1134, TITLE: Investigating the Impact of Multi-LiDAR Placement on Object Detection for Autonomous Driving
http://openaccess.thecvf.com//content/CVPR2022/html/Hu_Investigating_the_Impact_of_Multi-LiDAR_Placement_on_Object_Detection_for_CVPR_2022_paper.html
AUTHORS: Hanjiang Hu, Zuxin Liu, Sharad Chitlangia, Akhil Agnihotri, Ding Zhao
HIGHLIGHT: While most of the existing works focus on developing new deep learning algorithms or model architectures, we study the problem from the physical design perspective, i.e., how different placements of multiple LiDARs influence the learning-based perception. To this end, we introduce an easy-to-compute information-theoretic surrogate metric to quantitatively and fast evaluate LiDAR placement for 3D detection of different types of objects.

1135, TITLE: MViTv2: Improved Multiscale Vision Transformers for Classification and Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Li_MViTv2_Improved_Multiscale_Vision_Transformers_for_Classification_and_Detection_CVPR_2022_paper.html
AUTHORS: Yanghao Li, Chao-Yuan Wu, Haoqi Fan, Karttikeya Mangalam, Bo Xiong, Jitendra Malik, Christoph Feichtenhofer
HIGHLIGHT: In this paper, we study Multiscale Vision Transformers (MViTv2) as a unified architecture for image and video classification, as well as object detection.

1136, TITLE: Audio-Visual Generalised Zero-Shot Learning With Cross-Modal Attention and Language
http://openaccess.thecvf.com//content/CVPR2022/html/Mercea_Audio-Visual_Generalised_Zero-Shot_Learning_With_Cross-Modal_Attention_and_Language_CVPR_2022_paper.html
AUTHORS: Otniel-Bogdan Mercea, Lukas Riesch, A. Sophia Koepke, Zeynep Akata
HIGHLIGHT: Focusing on the relatively underexplored task of audio-visual zero-shot learning, we propose to learn multi-modal representations from audio-visual data using cross-modal attention and exploit textual label embeddings for transferring knowledge from seen classes to unseen classes.

1137, TITLE: Rethinking Efficient Lane Detection via Curve Modeling
http://openaccess.thecvf.com//content/CVPR2022/html/Feng_Rethinking_Efficient_Lane_Detection_via_Curve_Modeling_CVPR_2022_paper.html
AUTHORS: Zhengyang Feng, Shaohua Guo, Xin Tan, Ke Xu, Min Wang, Lizhuang Ma
HIGHLIGHT: This paper presents a novel parametric curve-based method for lane detection in RGB images.

1138, TITLE: GreedyNASv2: Greedier Search With a Greedy Path Filter
http://openaccess.thecvf.com//content/CVPR2022/html/Huang_GreedyNASv2_Greedier_Search_With_a_Greedy_Path_Filter_CVPR_2022_paper.html
AUTHORS: Tao Huang, Shan You, Fei Wang, Chen Qian, Changshui Zhang, Xiaogang Wang, Chang Xu
HIGHLIGHT: In this paper, we leverage an explicit path filter to capture the characteristics of paths and directly filter those weak ones, so that the search can be thus implemented on the shrunk space more greedily and efficiently.

1139, TITLE: Self-Supervised Arbitrary-Scale Point Clouds Upsampling via Implicit Neural Representation
http://openaccess.thecvf.com//content/CVPR2022/html/Zhao_Self-Supervised_Arbitrary-Scale_Point_Clouds_Upsampling_via_Implicit_Neural_Representation_CVPR_2022_paper.html
AUTHORS: Wenbo Zhao, Xianming Liu, Zhiwei Zhong, Junjun Jiang, Wei Gao, Ge Li, Xiangyang Ji
HIGHLIGHT: In this paper, we propose a novel approach that achieves selfsupervised and magnification-flexible point clouds upsampling simultaneously.

1140, TITLE: Co-Advise: Cross Inductive Bias Distillation
http://openaccess.thecvf.com//content/CVPR2022/html/Ren_Co-Advise_Cross_Inductive_Bias_Distillation_CVPR_2022_paper.html
AUTHORS: Sucheng Ren, Zhengqi Gao, Tianyu Hua, Zihui Xue, Yonglong Tian, Shengfeng He, Hang Zhao
HIGHLIGHT: Unlike previous works, where merely heavy convolution-based teachers are provided, in this paper, we delve into the influence of models inductive biases in knowledge distillation (e.g., convolution and involution).

1141, TITLE: AdaMixer: A Fast-Converging Query-Based Object Detector
http://openaccess.thecvf.com//content/CVPR2022/html/Gao_AdaMixer_A_Fast-Converging_Query-Based_Object_Detector_CVPR_2022_paper.html
AUTHORS: Ziteng Gao, Limin Wang, Bing Han, Sheng Guo
HIGHLIGHT: However, this paradigm still suffers from slow convergence, limited performance, and design complexity of extra networks between backbone and decoder. In this paper, we find that the key to these issues is the adaptability of decoders for casting queries to varying objects.

1142, TITLE: DTFD-MIL: Double-Tier Feature Distillation Multiple Instance Learning for Histopathology Whole Slide Image Classification
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_DTFD-MIL_Double-Tier_Feature_Distillation_Multiple_Instance_Learning_for_Histopathology_Whole_CVPR_2022_paper.html

AUTHORS: Hongrun Zhang, Yanda Meng, Yitian Zhao, Yihong Qiao, Xiaoyun Yang, Sarah E. Coupland, Yalin Zheng
HIGHLIGHT: In these, there are limited number of WSI slides (bags), while the resolution of a single WSI is huge, which leads to a large number of patches (instances) cropped from this slide. To address this issue, we propose to virtually enlarge the number of bags by introducing the concept of pseudo-bags, on which a double-tier MIL framework is built to effectively use the intrinsic features.

1143, TITLE: BEVT: BERT Pretraining of Video Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_BEVT_BERT_Pretraining_of_Video_Transformers_CVPR_2022_paper.html
AUTHORS: Rui Wang, Dongdong Chen, Zuxuan Wu, Yinpeng Chen, Xiyang Dai, Mengchen Liu, Yu-Gang Jiang, Luowei Zhou, Lu Yuan
HIGHLIGHT: We introduce BEVT which decouples video representation learning into spatial representation learning and temporal dynamics learning.

1144, TITLE: Deep Generalized Unfolding Networks for Image Restoration
http://openaccess.thecvf.com//content/CVPR2022/html/Mou_Deep_Generalized_Unfolding_Networks_for_Image_Restoration_CVPR_2022_paper.html
AUTHORS: Chong Mou, Qian Wang, Jian Zhang
HIGHLIGHT: In this paper, we propose a Deep Generalized Unfolding Network (DGUNet) for image restoration.

1145, TITLE: Automatic Relation-Aware Graph Network Proliferation
http://openaccess.thecvf.com//content/CVPR2022/html/Cai_Automatic_Relation-Aware_Graph_Network_Proliferation_CVPR_2022_paper.html
AUTHORS: Shaofei Cai, Liang Li, Xinzhe Han, Jiebo Luo, Zheng-Jun Zha, Qingming Huang
HIGHLIGHT: We propose Automatic Relation-aware Graph Network Proliferation (ARGNP) for efficiently searching GNNs with a relation-guided message passing mechanism.

1146, TITLE: AIM: An Auto-Augmenter for Images and Meshes
http://openaccess.thecvf.com//content/CVPR2022/html/Singh_AIM_An_Auto-Augmenter_for_Images_and_Meshes_CVPR_2022_paper.html
AUTHORS: Vinit Veerendraveer Singh, Chandra Kambhamettu
HIGHLIGHT: This paper presents an auto-augmenter for images and meshes (AIM) that easily incorporates into neural networks at training and inference times.

1147, TITLE: VISOLO: Grid-Based Space-Time Aggregation for Efficient Online Video Instance Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Han_VISOLO_Grid-Based_Space-Time_Aggregation_for_Efficient_Online_Video_Instance_Segmentation_CVPR_2022_paper.html
AUTHORS: Su Ho Han, Sukjun Hwang, Seoung Wug Oh, Yeonchool Park, Hyunwoo Kim, Min-Jung Kim, Seon Joo Kim
HIGHLIGHT: In this paper, we propose a novel single-stage framework for online VIS built based on the grid structured feature representation.

1148, TITLE: Deep Unlearning via Randomized Conditionally Independent Hessians
http://openaccess.thecvf.com//content/CVPR2022/html/Mehta_Deep_Unlearning_via_Randomized_Conditionally_Independent_Hessians_CVPR_2022_paper.html
AUTHORS: Ronak Mehta, Sourav Pal, Vikas Singh, Sathya N. Ravi
HIGHLIGHT: We use a variant of a new conditional independence coefficient, L-CODEC, to identify a subset of the model parameters with the most semantic overlap on an individual sample level.

1149, TITLE: Patch-Level Representation Learning for Self-Supervised Vision Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Yun_Patch-Level_Representation_Learning_for_Self-Supervised_Vision_Transformers_CVPR_2022_paper.html
AUTHORS: Sukmin Yun, Hankook Lee, Jaehyung Kim, Jinwoo Shin
HIGHLIGHT: Recent self-supervised learning (SSL) methods have shown impressive results in learning visual representations from unlabeled images. This paper aims to improve their performance further by utilizing the architectural advantages of the underlying neural network, as the current state-of-the-art visual pretext tasks for SSL do not enjoy the benefit, i.e., they are architecture-agnostic.

1150, TITLE: Sylph: A Hypernetwork Framework for Incremental Few-Shot Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Yin_Sylph_A_Hypernetwork_Framework_for_Incremental_Few-Shot_Object_Detection_CVPR_2022_paper.html
AUTHORS: Li Yin, Juan M. Perez-Rua, Kevin J. Liang
HIGHLIGHT: We study the challenging incremental few-shot object detection (iFSD) setting.

1151, TITLE: Incremental Learning in Semantic Segmentation From Image Labels
http://openaccess.thecvf.com//content/CVPR2022/html/Cermelli_Incremental_Learning_in_Semantic_Segmentation_From_Image_Labels_CVPR_2022_paper.html
AUTHORS: Fabio Cermelli, Dario Fontanel, Antonio Tavera, Marco Ciccone, Barbara Caputo
HIGHLIGHT: This paper proposes a novel framework for Weakly Incremental Learning for Semantic Segmentation, that aims at learning to segment new classes from cheap and largely available image-level labels.

1152, TITLE: Playable Environments: Video Manipulation in Space and Time
http://openaccess.thecvf.com/content/CVPR2022/html/Menapace_Playable_Environments_Video_Manipulation_in_Space_and_Time_CVPR_2022_paper.html
AUTHORS: Willi Menapace, Stéphane Lathuilière, Aliaksandr Siarohin, Christian Theobalt, Sergey Tulyakov, Vladislav Golyanik, Elisa Ricci
HIGHLIGHT: We present Playable Environments - a new representation for interactive video generation and manipulation in space and time.

1153, TITLE: Robust Cross-Modal Representation Learning With Progressive Self-Distillation
http://openaccess.thecvf.com//content/CVPR2022/html/Andonian_Robust_Cross-Modal_Representation_Learning_With_Progressive_Self-Distillation_CVPR_2022_paper.html
AUTHORS: Alex Andonian, Shixing Chen, Raffay Hamid
HIGHLIGHT: The learning objective of vision-language approach of CLIP does not effectively account for the noisy many-to-many correspondences found in web-harvested image captioning datasets, which contributes to its compute and data inefficiency. To address this challenge, we introduce a novel training framework based on cross-modal contrastive learning that uses progressive self-distillation and soft image-text alignments to more efficiently learn robust representations from noisy data.

1154, TITLE: What To Look at and Where: Semantic and Spatial Refined Transformer for Detecting Human-Object Interactions
http://openaccess.thecvf.com//content/CVPR2022/html/Iftekhar_What_To_Look_at_and_Where_Semantic_and_Spatial_Refined_CVPR_2022_paper.html
AUTHORS: A S M Iftekhar, Hao Chen, Kaustav Kundu, Xinyu Li, Joseph Tighe, Davide Modolo
HIGHLIGHT: We propose a novel one-stage Transformer-based semantic and spatial refined transformer (SSRT) to solve the Human-Object Interaction detection task, which requires to localize humans and objects, and predicts their interactions.

1155, TITLE: Compressive Single-Photon 3D Cameras
http://openaccess.thecvf.com//content/CVPR2022/html/Gutierrez-Barragan_Compressive_Single-Photon_3D_Cameras_CVPR_2022_paper.html
AUTHORS: Felipe Gutierrez-Barragan, Atul Ingle, Trevor Seets, Mohit Gupta, Andreas Velten
HIGHLIGHT: One major reason for SPAD's bandwidth-intensive operation is the tight coupling that exists between depth resolution and histogram resolution. To weaken this coupling, we propose compressive single-photon histograms (CSPH).

1156, TITLE: Stereo Magnification With Multi-Layer Images
http://openaccess.thecvf.com//content/CVPR2022/html/Khakhulin_Stereo_Magnification_With_Multi-Layer_Images_CVPR_2022_paper.html
AUTHORS: Taras Khakhulin, Denis Korzhenkov, Pavel Solovev, Gleb Sterkin, Andrei-Timotei Ardelean, Victor Lempitsky
HIGHLIGHT: In this work, we introduce a new view synthesis approach based on multiple semitransparent layers with scene-adapted geometry.

1157, TITLE: CO-SNE: Dimensionality Reduction and Visualization for Hyperbolic Data
http://openaccess.thecvf.com//content/CVPR2022/html/Guo_CO-SNE_Dimensionality_Reduction_and_Visualization_for_Hyperbolic_Data_CVPR_2022_paper.html
AUTHORS: Yunhui Guo, Haoran Guo, Stella X. Yu
HIGHLIGHT: We propose CO-SNE, which extends the Euclidean space visualization tool, t-SNE, to hyperbolic space.

1158, TITLE: Revisiting Skeleton-Based Action Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Duan_Revisiting_Skeleton-Based_Action_Recognition_CVPR_2022_paper.html
AUTHORS: Haodong Duan, Yue Zhao, Kai Chen, Dahua Lin, Bo Dai
HIGHLIGHT: In this work, we propose PoseConv3D, a new approach to skeleton-based action recognition.

1159, TITLE: Rethinking Controllable Variational Autoencoders
http://openaccess.thecvf.com//content/CVPR2022/html/Shao_Rethinking_Controllable_Variational_Autoencoders_CVPR_2022_paper.html
AUTHORS: Huajie Shao, Yifei Yang, Haohong Lin, Longzhong Lin, Yizhuo Chen, Qinmin Yang, Han Zhao
HIGHLIGHT: However, when it comes to disentangled representation learning, ControlVAE does not delve into the rationale behind it. The goal of this paper is to develop a deeper understanding of ControlVAE in learning disentangled representations, including the choice of a desired KL-divergence (i.e, set point), and its stability during training.

1160, TITLE: Contextual Instance Decoupling for Robust Multi-Person Pose Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Contextual_Instance_Decoupling_for_Robust_Multi-Person_Pose_Estimation_CVPR_2022_paper.html
AUTHORS: Dongkai Wang, Shiliang Zhang
HIGHLIGHT: This paper proposes the Contextual Instance Decoupling (CID), which presents a new pipeline for multi-person pose estimation.

1161, TITLE: LMGP: Lifted Multicut Meets Geometry Projections for Multi-Camera Multi-Object Tracking
http://openaccess.thecvf.com//content/CVPR2022/html/Nguyen_LMGP_Lifted_Multicut_Meets_Geometry_Projections_for_Multi-Camera_Multi-Object_Tracking_CVPR_2022_paper.html
AUTHORS: Duy M. H. Nguyen, Roberto Henschel, Bodo Rosenhahn, Daniel Sonntag, Paul Swoboda

HIGHLIGHT:	In this work, we propose a mathematically elegant multi-camera multiple object tracking approach based on a spatial-temporal lifted multicut formulation.

1162, TITLE:	Boosting Crowd Counting via Multifaceted Attention
http://openaccess.thecvf.com//content/CVPR2022/html/Lin_Boosting_Crowd_Counting_via_Multifaceted_Attention_CVPR_2022_paper.html
AUTHORS:	Hui Lin, Zhiheng Ma, Rongrong Ji, Yaowei Wang, Xiaopeng Hong
HIGHLIGHT:	As large-scale variations often exist within crowd images, neither fixed-size convolution kernel of CNN nor fixed-size attentions of recent vision transformers can well handle this kind of variations. To address this problem, we propose a Multifaceted Attention Network (MAN), which incorporates global attention from vanilla transformer, learnable local attention, attention regularization and instance attention into a counting model.

1163, TITLE:	Stereo Depth From Events Cameras: Concentrate and Focus on the Future
http://openaccess.thecvf.com//content/CVPR2022/html/Nam_Stereo_Depth_From_Events_Cameras_Concentrate_and_Focus_on_the_CVPR_2022_paper.html
AUTHORS:	Yeongwoo Nam, Mohammad Mostafavi, Kuk-Jin Yoon, Jonghyun Choi
HIGHLIGHT:	To alleviate the event missing or overriding issue, we propose to learn to concentrate on the dense events to produce a compact event representation with high details for depth estimation.

1164, TITLE:	A Probabilistic Graphical Model Based on Neural-Symbolic Reasoning for Visual Relationship Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Yu_A_Probabilistic_Graphical_Model_Based_on_Neural-Symbolic_Reasoning_for_Visual_CVPR_2022_paper.html
AUTHORS:	Dongran Yu, Bo Yang, Qianhao Wei, Anchen Li, Shirui Pan
HIGHLIGHT:	To overcome the aforementioned weaknesses, we integrate symbolic knowledge into deep learning models and propose a bi-level probabilistic graphical reasoning framework called BPGR.

1165, TITLE:	A Simple Data Mixing Prior for Improving Self-Supervised Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Ren_A_Simple_Data_Mixing_Prior_for_Improving_Self-Supervised_Learning_CVPR_2022_paper.html
AUTHORS:	Sucheng Ren, Huiyu Wang, Zhengqi Gao, Shengfeng He, Alan Yuille, Yuyin Zhou, Cihang Xie
HIGHLIGHT:	Data mixing (e.g., Mixup, Cutmix, ResizeMix) is an essential component for advancing recognition models. In this paper, we focus on studying its effectiveness in the self-supervised setting.

1166, TITLE:	Knowledge Distillation As Efficient Pre-Training: Faster Convergence, Higher Data-Efficiency, and Better Transferability
http://openaccess.thecvf.com//content/CVPR2022/html/He_Knowledge_Distillation_As_Efficient_Pre-Training_Faster_Convergence_Higher_Data-Efficiency_and_CVPR_2022_paper.html
AUTHORS:	Ruifei He, Shuyang Sun, Jihan Yang, Song Bai, Xiaojuan Qi
HIGHLIGHT:	In this work, we investigate an alternative strategy for pre-training, namely Knowledge Distillation as Efficient Pre-training (KDEP), aiming to efficiently transfer the learned feature representation from existing pre-trained models to new student models for future downstream tasks.

1167, TITLE:	LOLNerf: Learn From One Look
http://openaccess.thecvf.com//content/CVPR2022/html/Rebain_LOLNerf_Learn_From_One_Look_CVPR_2022_paper.html
AUTHORS:	Daniel Rebain, Mark Matthews, Kwang Moo Yi, Dmitry Lagun, Andrea Tagliasacchi
HIGHLIGHT:	We present a method for learning a generative 3D model based on neural radiance fields, trained solely from data with only single views of each object.

1168, TITLE:	Geometry-Aware Guided Loss for Deep Crack Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Geometry-Aware_Guided_Loss_for_Deep_Crack_Recognition_CVPR_2022_paper.html
AUTHORS:	Zhuangzhuang Chen, Jin Zhang, Zhuonan Lai, Jie Chen, Zun Liu, Jianqiang Li
HIGHLIGHT:	In this paper, we propose the geometry-aware guided loss (GAGL) that enhances the discrimination ability and is only applied in the training stage without extra computation and memory during inference.

1169, TITLE:	Multi-Modal Alignment Using Representation Codebook
http://openaccess.thecvf.com//content/CVPR2022/html/Duan_Multi-Modal_Alignment_Using_Representation_Codebook_CVPR_2022_paper.html
AUTHORS:	Jiali Duan, Liqun Chen, Son Tran, Jinyu Yang, Yi Xu, Belinda Zeng, Trishul Chilimbi
HIGHLIGHT:	In this paper, we propose to align at a higher and more stable level using cluster representation.

1170, TITLE:	Maintaining Reasoning Consistency in Compositional Visual Question Answering
http://openaccess.thecvf.com//content/CVPR2022/html/Jing_Maintaining_Reasoning_Consistency_in_Compositional_Visual_Question_Answering_CVPR_2022_paper.html
AUTHORS:	Chenchen Jing, Yunde Jia, Yuwei Wu, Xinyu Liu, Qi Wu
HIGHLIGHT:	This paper presents a dialog-like reasoning method for maintaining reasoning consistency in answering a compositional question and its sub-questions.

1171, TITLE:	Structure-Aware Motion Transfer With Deformable Anchor Model

http://openaccess.thecvf.com//content/CVPR2022/html/Tao_Structure-Aware_Motion_Transfer_With_Deformable_Anchor_Model_CVPR_2022_paper.html
AUTHORS: Jiale Tao, Biao Wang, Borun Xu, Tiezheng Ge, Yuning Jiang, Wen Li, Lixin Duan
HIGHLIGHT: In this paper, we propose a novel structure-aware motion modeling approach, the deformable anchor model (DAM), which can automatically discover the motion structure of arbitrary objects without leveraging their prior structure information.

1172, TITLE: BigDL 2.0: Seamless Scaling of AI Pipelines From Laptops to Distributed Cluster
http://openaccess.thecvf.com//content/CVPR2022/html/Dai_BigDL_2.0_Seamless_Scaling_of_AI_Pipelines_From_Laptops_to_CVPR_2022_paper.html
AUTHORS: Jason (Jinquan) Dai, Ding Ding, Dongjie Shi, Shengsheng Huang, Jiao Wang, Xin Qiu, Kai Huang, Guoqiong Song, Yang Wang, Qiyuan Gong, Jiaming Song, Shan Yu, Le Zheng, Yina Chen, Junwei Deng, Ge Song
HIGHLIGHT: To address this challenge, we have open sourced BigDL 2.0 at https://github.com/intel-analytics/BigDL/ under Apache 2.0 license (combining the original BigDL [19] and Analytics Zoo [18] projects); using BigDL 2.0, users can simply build conventional Python notebooks on their laptops (with possible AutoML support), which can then be transparently accelerated on a single node (with up-to 9.6x speedup in our experiments), and seamlessly scaled out to a large cluster (across several hundreds servers in real-world use cases).

1173, TITLE: Integrative Few-Shot Learning for Classification and Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Kang_Integrative_Few-Shot_Learning_for_Classification_and_Segmentation_CVPR_2022_paper.html
AUTHORS: Dahyun Kang, Minsu Cho
HIGHLIGHT: We introduce the integrative task of few-shot classification and segmentation (FS-CS) that aims to both classify and segment target objects in a query image when the target classes are given with a few examples.

1174, TITLE: Acquiring a Dynamic Light Field Through a Single-Shot Coded Image
http://openaccess.thecvf.com//content/CVPR2022/html/Mizuno_Acquiring_a_Dynamic_Light_Field_Through_a_Single-Shot_Coded_Image_CVPR_2022_paper.html
AUTHORS: Ryoya Mizuno, Keita Takahashi, Michitaka Yoshida, Chihiro Tsutake, Toshiaki Fujii, Hajime Nagahara
HIGHLIGHT: We propose a method for compressively acquiring a dynamic light field (a 5-D volume) through a single-shot coded image (a 2-D measurement).

1175, TITLE: Attentive Fine-Grained Structured Sparsity for Image Restoration
http://openaccess.thecvf.com//content/CVPR2022/html/Oh_Attentive_Fine-Grained_Structured_Sparsity_for_Image_Restoration_CVPR_2022_paper.html
AUTHORS: Junghun Oh, Heewon Kim, Seungjun Nah, Cheeun Hong, Jonghyun Choi, Kyoung Mu Lee
HIGHLIGHT: To further optimize the trade-off between the efficiency and the restoration accuracy, we propose a novel pruning method that determines the pruning ratio for N:M structured sparsity at each layer.

1176, TITLE: Pix2NeRF: Unsupervised Conditional p-GAN for Single Image to Neural Radiance Fields Translation
http://openaccess.thecvf.com//content/CVPR2022/html/Cai_Pix2NeRF_Unsupervised_Conditional_p-GAN_for_Single_Image_to_Neural_Radiance_CVPR_2022_paper.html
AUTHORS: Shengqu Cai, Anton Obukhov, Dengxin Dai, Luc Van Gool
HIGHLIGHT: We propose a pipeline to generate Neural Radiance Fields (NeRF) of an object or a scene of a specific class, conditioned on a single input image.

1177, TITLE: HARA: A Hierarchical Approach for Robust Rotation Averaging
http://openaccess.thecvf.com//content/CVPR2022/html/Lee_HARA_A_Hierarchical_Approach_for_Robust_Rotation_Averaging_CVPR_2022_paper.html
AUTHORS: Seong Hun Lee, Javier Civera
HIGHLIGHT: We propose a novel hierarchical approach for multiple rotation averaging, dubbed HARA.

1178, TITLE: Diffusion Autoencoders: Toward a Meaningful and Decodable Representation
http://openaccess.thecvf.com//content/CVPR2022/html/Preechakul_Diffusion_Autoencoders_Toward_a_Meaningful_and_Decodable_Representation_CVPR_2022_paper.html
AUTHORS: Konpat Preechakul, Nattanat Chatthee, Suttisak Wizadwongsa, Supasorn Suwajanakorn
HIGHLIGHT: Our key idea is to use a learnable encoder for discovering the high-level semantics, and a DPM as the decoder for modeling the remaining stochastic variations.

1179, TITLE: Learning Fair Classifiers With Partially Annotated Group Labels
http://openaccess.thecvf.com//content/CVPR2022/html/Jung_Learning_Fair_Classifiers_With_Partially_Annotated_Group_Labels_CVPR_2022_paper.html
AUTHORS: Sangwon Jung, Sanghyuk Chun, Taesup Moon
HIGHLIGHT: In this paper, we consider a more practical scenario, dubbed as Algorithmic Group Fairness with the Partially annotated Group labels (Fair-PG).

1180, TITLE: StylizedNeRF: Consistent 3D Scene Stylization As Stylized NeRF via 2D-3D Mutual Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Huang_StylizedNeRF_Consistent_3D_Scene_Stylization_As_Stylized_NeRF_via_2D-3D_CVPR_2022_paper.html
AUTHORS: Yi-Hua Huang, Yue He, Yu-Jie Yuan, Yu-Kun Lai, Lin Gao

HIGHLIGHT: However, there is a significant domain gap between style examples which are 2D images and NeRF which is an implicit volumetric representation. To address this problem, we propose a novel mutual learning framework for 3D scene stylization that combines a 2D image stylization network and NeRF to fuse the stylization ability of 2D stylization network with the 3D consistency of NeRF.

1181, TITLE: NightLab: A Dual-Level Architecture With Hardness Detection for Segmentation at Night
http://openaccess.thecvf.com//content/CVPR2022/html/Deng_NightLab_A_Dual-
Level_Architecture_With_Hardness_Detection_for_Segmentation_at_CVPR_2022_paper.html
AUTHORS: Xueqing Deng, Peng Wang, Xiaochen Lian, Shawn Newsam
HIGHLIGHT: In this paper, we propose NightLab, a novel nighttime segmentation framework that leverages multiple deep learning models imbued with night-aware features to yield State-of-The-Art (SoTA) performance on multiple night segmentation benchmarks.

1182, TITLE: Knowledge Distillation With the Reused Teacher Classifier
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Knowledge_Distillation_With_the_Reused_Teacher_Classifier_CVPR_2022_paper.html
AUTHORS: Defang Chen, Jian-Ping Mei, Hailin Zhang, Can Wang, Yan Feng, Chun Chen
HIGHLIGHT: In contrast, we empirically show that a simple knowledge distillation technique is enough to significantly narrow down the teacher-student performance gap.

1183, TITLE: Contrastive Learning for Unsupervised Video Highlight Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Badamdorj_Contrastive_Learning_for_Unsupervised_Video_Highlight_Detection_CVPR_2022_paper.html
AUTHORS: Taivanbat Badamdorj, Mrigank Rochan, Yang Wang, Li Cheng
HIGHLIGHT: We propose a simple contrastive learning framework for unsupervised highlight detection.

1184, TITLE: InfoGCN: Representation Learning for Human Skeleton-Based Action Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Chi_InfoGCN_Representation_Learning_for_Human_Skeleton-
Based_Action_Recognition_CVPR_2022_paper.html
AUTHORS: Hyung-gun Chi, Myoung Hoon Ha, Seunggeun Chi, Sang Wan Lee, Qixing Huang, Karthik Ramani
HIGHLIGHT: InfoGCN proposes a learning framework for action recognition combining a novel learning objective and an encoding method.

1185, TITLE: Rethinking Image Cropping: Exploring Diverse Compositions From Global Views
http://openaccess.thecvf.com//content/CVPR2022/html/Jia_Rethinking_Image_Cropping_Exploring_Diverse_Compositions_From_Global_Views_CVPR_2022_paper.html
AUTHORS: Gengyun Jia, Huaibo Huang, Chaoyou Fu, Ran He
HIGHLIGHT: In this paper, we regard image cropping as a set prediction problem.

1186, TITLE: Constrained Few-Shot Class-Incremental Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Hersche_Constrained_Few-Shot_Class-
Incremental_Learning_CVPR_2022_paper.html
AUTHORS: Michael Hersche, Geethan Karunaratne, Giovanni Cherubini, Luca Benini, Abu Sebastian, Abbas Rahimi
HIGHLIGHT: To meet the above constraints, we propose C-FSCIL, which is architecturally composed of a frozen meta-learned feature extractor, a trainable fixed-size fully connected layer, and a rewritable dynamically growing memory that stores as many vectors as the number of encountered classes.

1187, TITLE: Self-Supervised Material and Texture Representation Learning for Remote Sensing Tasks
http://openaccess.thecvf.com//content/CVPR2022/html/Akiva_Self-
Supervised_Material_and_Texture_Representation_Learning_for_Remote_Sensing_Tasks_CVPR_2022_paper.html
AUTHORS: Peri Akiva, Matthew Purri, Matthew Leotta
HIGHLIGHT: In this work, we present our material and texture based self-supervision method named MATTER (MATerial and TExture Representation Learning), which is inspired by classical material and texture methods.

1188, TITLE: Threshold Matters in WSSS: Manipulating the Activation for the Robust and Accurate Segmentation Model Against Thresholds
http://openaccess.thecvf.com//content/CVPR2022/html/Lee_Threshold_Matters_in_WSSS_Manipulating_the_Activation_for_the_Robust_CVPR_2022_paper.html
AUTHORS: Minhyun Lee, Dongseob Kim, Hyunjung Shim
HIGHLIGHT: Then, we show that this issue can be mitigated by satisfying two conditions; 1) reducing the imbalance in the foreground activation and 2) increasing the gap between the foreground and the background activation. Based on these findings, we propose a novel activation manipulation network with a per-pixel classification loss and a label conditioning module.

1189, TITLE: Data-Free Network Compression via Parametric Non-Uniform Mixed Precision Quantization
http://openaccess.thecvf.com//content/CVPR2022/html/Chikin_Data-Free_Network_Compression_via_Parametric_Non-
Uniform_Mixed_Precision_Quantization_CVPR_2022_paper.html
AUTHORS: Vladimir Chikin, Mikhail Antiukh
HIGHLIGHT: These are unavailable for the confidential scenarios due to personal privacy and security problems. Focusing on this issue, we propose a novel data-free method for network compression called PNMQ, which employs the Parametric Non-uniform Mixed precision Quantization to generate a quantized network.

126

1190, TITLE: Sparse to Dense Dynamic 3D Facial Expression Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Otberdout_Sparse_to_Dense_Dynamic_3D_Facial_Expression_Generation_CVPR_2022_paper.html
AUTHORS: Naima Otberdout, Claudio Ferrari, Mohamed Daoudi, Stefano Berretti, Alberto Del Bimbo
HIGHLIGHT: In this paper, we propose a solution to the task of generating dynamic 3D facial expressions from a neutral 3D face and an expression label.

1191, TITLE: Think Twice Before Detecting GAN-Generated Fake Images From Their Spectral Domain Imprints
http://openaccess.thecvf.com//content/CVPR2022/html/Dong_Think_Twice_Before_Detecting_GAN-Generated_Fake_Images_From_Their_Spectral_CVPR_2022_paper.html
AUTHORS: Chengdong Dong, Ajay Kumar, Eryun Liu
HIGHLIGHT: Our study in this paper introduces a pipeline to mitigate the spectral artifacts.

1192, TITLE: Crafting Better Contrastive Views for Siamese Representation Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Peng_Crafting_Better_Contrastive_Views_for_Siamese_Representation_Learning_CVPR_2022_paper.html
AUTHORS: Xiangyu Peng, Kai Wang, Zheng Zhu, Mang Wang, Yang You
HIGHLIGHT: In this work, we propose ContrastiveCrop, which could effectively generate better crops for Siamese representation learning.

1193, TITLE: RSCFed: Random Sampling Consensus Federated Semi-Supervised Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Liang_RSCFed_Random_Sampling_Consensus_Federated_Semi-Supervised_Learning_CVPR_2022_paper.html
AUTHORS: Xiaoxiao Liang, Yiqun Lin, Huazhu Fu, Lei Zhu, Xiaomeng Li
HIGHLIGHT: In this paper, we present a Random Sampling Consensus Federated learning, namely RSCFed, by considering the uneven reliability among models from labeled clients and unlabeled clients.

1194, TITLE: TransMVSNet: Global Context-Aware Multi-View Stereo Network With Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Ding_TransMVSNet_Global_Context-Aware_Multi-View_Stereo_Network_With_Transformers_CVPR_2022_paper.html
AUTHORS: Yikang Ding, Wentao Yuan, Qingtian Zhu, Haotian Zhang, Xiangyue Liu, Yuanjiang Wang, Xiao Liu
HIGHLIGHT: In this paper, we present TransMVSNet, based on our exploration of feature matching in multi-view stereo (MVS).

1195, TITLE: ROCA: Robust CAD Model Retrieval and Alignment From a Single Image
http://openaccess.thecvf.com//content/CVPR2022/html/Gumeli_ROCA_Robust_CAD_Model_Retrieval_and_Alignment_From_a_Single_CVPR_2022_paper.html
AUTHORS: Can Gümeli, Angela Dai, Matthias Nießner
HIGHLIGHT: We present ROCA, a novel end-to-end approach that retrieves and aligns 3D CAD models from a shape database to a single input image.

1196, TITLE: Continual Learning for Visual Search With Backward Consistent Feature Embedding
http://openaccess.thecvf.com//content/CVPR2022/html/Wan_Continual_Learning_for_Visual_Search_With_Backward_Consistent_Feature_Embedding_CVPR_2022_paper.html
AUTHORS: Timmy S. T. Wan, Jun-Cheng Chen, Tzer-Yi Wu, Chu-Song Chen
HIGHLIGHT: To address the issues of long-term visual search, we introduce a continual learning (CL) approach that can handle the incrementally growing gallery set with backward embedding consistency.

1197, TITLE: iFS-RCNN: An Incremental Few-Shot Instance Segmenter
http://openaccess.thecvf.com//content/CVPR2022/html/Nguyen_iFS-RCNN_An_Incremental_Few-Shot_Instance_Segmenter_CVPR_2022_paper.html
AUTHORS: Khoi Nguyen, Sinisa Todorovic
HIGHLIGHT: This paper addresses incremental few-shot instance segmentation, where a few examples of new object classes arrive when access to training examples of old classes is not available anymore, and the goal is to perform well on both old and new classes.

1198, TITLE: DPGEN: Differentially Private Generative Energy-Guided Network for Natural Image Synthesis
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_DPGEN_Differentially_Private_Generative_Energy-Guided_Network_for_Natural_Image_Synthesis_CVPR_2022_paper.html
AUTHORS: Jia-Wei Chen, Chia-Mu Yu, Ching-Chia Kao, Tzai-Wei Pang, Chun-Shien Lu
HIGHLIGHT: Here, we propose DPGEN, a network model designed to synthesize high-resolution natural images while satisfying differential privacy.

1199, TITLE: MetaFSCIL: A Meta-Learning Approach for Few-Shot Class Incremental Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Chi_MetaFSCIL_A_Meta-Learning_Approach_for_Few-Shot_Class_Incremental_Learning_CVPR_2022_paper.html
AUTHORS: Zhixiang Chi, Li Gu, Huan Liu, Yang Wang, Yuanhao Yu, Jin Tang
HIGHLIGHT: In this paper, we tackle the problem of few-shot class incremental learning (FSCIL).

1200, TITLE: The Majority Can Help the Minority: Context-Rich Minority Oversampling for Long-Tailed Classification
http://openaccess.thecvf.com//content/CVPR2022/html/Park_The_Majority_Can_Help_the_Minority_Context-Rich_Minority_Oversampling_for_CVPR_2022_paper.html
AUTHORS: Seulki Park, Youngkyu Hong, Byeongho Heo, Sangdoo Yun, Jin Young Choi
HIGHLIGHT: In this paper, we propose a novel minority over-sampling method to augment diversified minority samples by leveraging the rich context of the majority classes as background images.

1201, TITLE: Dense Depth Priors for Neural Radiance Fields From Sparse Input Views
http://openaccess.thecvf.com//content/CVPR2022/html/Roessle_Dense_Depth_Priors_for_Neural_Radiance_Fields_From_Sparse_Input_CVPR_2022_paper.html
AUTHORS: Barbara Roessle, Jonathan T. Barron, Ben Mildenhall, Pratul P. Srinivasan, Matthias Nießner
HIGHLIGHT: Our method aims to synthesize novel views of whole rooms from an order of magnitude fewer images.

1202, TITLE: EyePAD++: A Distillation-Based Approach for Joint Eye Authentication and Presentation Attack Detection Using Periocular Images
http://openaccess.thecvf.com//content/CVPR2022/html/Dhar_EyePAD_A_Distillation-Based_Approach_for_Joint_Eye_Authentication_and_Presentation_CVPR_2022_paper.html
AUTHORS: Prithviraj Dhar, Amit Kumar, Kirsten Kaplan, Khushi Gupta, Rakesh Ranjan, Rama Chellappa
HIGHLIGHT: Here, we introduce a joint framework for EA and PAD using periocular images.

1203, TITLE: IntentVizor: Towards Generic Query Guided Interactive Video Summarization
http://openaccess.thecvf.com//content/CVPR2022/html/Wu_IntentVizor_Towards_Generic_Query_Guided_Interactive_Video_Summarization_CVPR_2022_paper.html
AUTHORS: Guande Wu, Jianzhe Lin, Claudio T. Silva
HIGHLIGHT: First, the text query might not be enough to describe the exact and diverse needs of the user. Second, the user cannot edit once the summaries are produced, while we assume the needs of the user should be subtle and need to be adjusted interactively. To solve these two problems, we propose IntentVizor, an interactive video summarization framework guided by generic multi-modality queries.

1204, TITLE: Wnet: Audio-Guided Video Object Segmentation via Wavelet-Based Cross-Modal Denoising Networks
http://openaccess.thecvf.com//content/CVPR2022/html/Pan_Wnet_Audio-Guided_Video_Object_Segmentation_via_Wavelet-Based_Cross-Modal_Denoising_Networks_CVPR_2022_paper.html
AUTHORS: Wenwen Pan, Haonan Shi, Zhou Zhao, Jieming Zhu, Xiuqiang He, Zhigeng Pan, Lianli Gao, Jun Yu, Fei Wu, Qi Tian
HIGHLIGHT: In this paper, we consider the problem of audio-guided video semantic segmentation from the viewpoint of end-to-end denoised encoder-decoder network learning.

1205, TITLE: Camera Pose Estimation Using Implicit Distortion Models
http://openaccess.thecvf.com//content/CVPR2022/html/Pan_Camera_Pose_Estimation_Using_Implicit_Distortion_Models_CVPR_2022_paper.html
AUTHORS: Linfei Pan, Marc Pollefeys, Viktor Larsson
HIGHLIGHT: In this paper, we explore an alternative approach which implicitly models the lens distortion.

1206, TITLE: Scene Representation Transformer: Geometry-Free Novel View Synthesis Through Set-Latent Scene Representations
http://openaccess.thecvf.com//content/CVPR2022/html/Sajjadi_Scene_Representation_Transformer_Geometry-Free_Novel_View_Synthesis_Through_Set-Latent_Scene_CVPR_2022_paper.html
AUTHORS: Mehdi S. M. Sajjadi, Henning Meyer, Etienne Pot, Urs Bergmann, Klaus Greff, Noha Radwan, Suhani Vora, Mario LuÄ iÄ‡, Daniel Duckworth, Alexey Dosovitskiy, Jakob Uszkoreit, Thomas Funkhouser, Andrea Tagliasacchi
HIGHLIGHT: In this work, we propose the Scene Representation Transformer (SRT), a method which processes posed or unposed RGB images of a new area, infers a "set-latent scene representation", and synthesises novel views, all in a single feed-forward pass.

1207, TITLE: Shape-Invariant 3D Adversarial Point Clouds
http://openaccess.thecvf.com//content/CVPR2022/html/Huang_Shape-Invariant_3D_Adversarial_Point_Clouds_CVPR_2022_paper.html
AUTHORS: Qidong Huang, Xiaoyi Dong, Dongdong Chen, Hang Zhou, Weiming Zhang, Nenghai Yu
HIGHLIGHT: In this paper, we propose a novel Point-Cloud Sensitivity Map to boost both the efficiency and imperceptibility of point perturbations.

1208, TITLE: LAS-AT: Adversarial Training With Learnable Attack Strategy
http://openaccess.thecvf.com//content/CVPR2022/html/Jia_LAS-AT_Adversarial_Training_With_Learnable_Attack_Strategy_CVPR_2022_paper.html
AUTHORS: Xiaojun Jia, Yong Zhang, Baoyuan Wu, Ke Ma, Jue Wang, Xiaochun Cao
HIGHLIGHT: In this paper, we propose a novel framework for adversarial training by introducing the concept of "learnable attack strategy", dubbed LAS-AT, which learns to automatically produce attack strategies to improve the model robustness.

1209, TITLE: Bootstrapping ViTs: Towards Liberating Vision Transformers From Pre-Training

http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Bootstrapping_ViTs_Towards_Liberating_Vision_Transformers_From_Pre-Training_CVPR_2022_paper.html
AUTHORS:	Haofei Zhang, Jiarui Duan, Mengqi Xue, Jie Song, Li Sun, Mingli Song
HIGHLIGHT:	In this paper, we strive to liberate ViTs from pre-training by introducing CNNs' inductive biases back to ViTs while preserving their network architectures for higher upper bound and setting up more suitable optimization objectives.

1210, TITLE:	PubTables-1M: Towards Comprehensive Table Extraction From Unstructured Documents
http://openaccess.thecvf.com//content/CVPR2022/html/Smock_PubTables-1M_Towards_Comprehensive_Table_Extraction_From_Unstructured_Documents_CVPR_2022_paper.html
AUTHORS:	Brandon Smock, Rohith Pesala, Robin Abraham
HIGHLIGHT:	However, one of the greatest challenges remains the creation of datasets with complete, unambiguous ground truth at scale. To address this, we develop a new, more comprehensive dataset for table extraction, called PubTables-1M.

1211, TITLE:	Styleformer: Transformer Based Generative Adversarial Networks With Style Vector
http://openaccess.thecvf.com//content/CVPR2022/html/Park_Styleformer_Transformer_Based_Generative_Adversarial_Networks_With_Style_Vector_CVPR_2022_paper.html
AUTHORS:	Jeeseung Park, Younggeun Kim
HIGHLIGHT:	In this paper, we effectively apply the modified Transformer structure (e.g., Increased multi-head attention and Pre-layer normalization) and attention style injection which is style modulation and demodulation method for self-attention operation.

1212, TITLE:	Efficient Two-Stage Detection of Human-Object Interactions With a Novel Unary-Pairwise Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Efficient_Two-Stage_Detection_of_Human-Object_Interactions_With_a_Novel_Unary-Pairwise_CVPR_2022_paper.html
AUTHORS:	Frederic Z. Zhang, Dylan Campbell, Stephen Gould
HIGHLIGHT:	In this work, we propose the Unary-Pairwise Transformer, a two-stage detector that exploits unary and pairwise representations for HOIs.

1213, TITLE:	ELSR: Efficient Line Segment Reconstruction With Planes and Points Guidance
http://openaccess.thecvf.com//content/CVPR2022/html/Wei_ELSR_Efficient_Line_Segment_Reconstruction_With_Planes_and_Points_Guidance_CVPR_2022_paper.html
AUTHORS:	Dong Wei, Yi Wan, Yongjun Zhang, Xinyi Liu, Bin Zhang, Xiqi Wang
HIGHLIGHT:	In this paper, we propose an efficient line segment reconstruction method called ELSR.

1214, TITLE:	Meta-Attention for ViT-Backed Continual Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Xue_Meta-Attention_for_ViT-Backed_Continual_Learning_CVPR_2022_paper.html
AUTHORS:	Mengqi Xue, Haofei Zhang, Jie Song, Mingli Song
HIGHLIGHT:	In this paper, we study ViT-backed continual learning to strive for higher performance riding on recent advances of ViTs.

1215, TITLE:	DST: Dynamic Substitute Training for Data-Free Black-Box Attack
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_DST_Dynamic_Substitute_Training_for_Data-Free_Black-Box_Attack_CVPR_2022_paper.html
AUTHORS:	Wenxuan Wang, Xuelin Qian, Yanwei Fu, Xiangyang Xue
HIGHLIGHT:	In this paper, we propose a novel dynamic substitute training attack method to encourage substitute model to learn better and faster from the target model.

1216, TITLE:	Photorealistic Monocular 3D Reconstruction of Humans Wearing Clothing
http://openaccess.thecvf.com//content/CVPR2022/html/Alldieck_Photorealistic_Monocular_3D_Reconstruction_of_Humans_Wearing_Clothing_CVPR_2022_paper.html
AUTHORS:	Thiemo Alldieck, Mihai Zanfir, Cristian Sminchisescu
HIGHLIGHT:	We present PHORHUM, a novel, end-to-end trainable, deep neural network methodology for photorealistic 3D human reconstruction given just a monocular RGB image.

1217, TITLE:	A Low-Cost & Real-Time Motion Capture System
http://openaccess.thecvf.com//content/CVPR2022/html/Chatzitofis_A_Low-Cost__Real-Time_Motion_Capture_System_CVPR_2022_paper.html
AUTHORS:	Anargyros Chatzitofis, Georgios Albanis, Nikolaos Zioulis, Spyridon Thermos
HIGHLIGHT:	Traditional marker-based motion capture requires excessive and specialized equipment, hindering accessibility and wider adoption. In this work, we demonstrate such a system but rely on a very sparse set of low-cost consumer-grade sensors.

1218, TITLE:	Unified Contrastive Learning in Image-Text-Label Space
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Unified_Contrastive_Learning_in_Image-Text-Label_Space_CVPR_2022_paper.html
AUTHORS:	Jianwei Yang, Chunyuan Li, Pengchuan Zhang, Bin Xiao, Ce Liu, Lu Yuan, Jianfeng Gao
HIGHLIGHT:	In this work, we introduce a new formulation by combining the two data sources into a common image-text-label space.

1219, TITLE:	Unifying Motion Deblurring and Frame Interpolation With Events

http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Unifying_Motion_Deblurring_and_Frame_Interpolation_With_Events_CVPR_2022_paper.html

AUTHORS: Xiang Zhang, Lei Yu

HIGHLIGHT: Slow shutter speed and long exposure time of frame-based cameras often cause visual blur and loss of inter-frame information, degenerating the overall quality of captured videos. To this end, we present a unified framework of event-based motion deblurring and frame interpolation for blurry video enhancement, where the extremely low latency of events is leveraged to alleviate motion blur and facilitate intermediate frame prediction.

1220, TITLE: Generalizing Interactive Backpropagating Refinement for Dense Prediction Networks

http://openaccess.thecvf.com//content/CVPR2022/html/Lin_Generalizing_Interactive_Backpropagating_Refinement_for_Dense_Prediction_Networks_CVPR_2022_paper.html

AUTHORS: Fanqing Lin, Brian Price, Tony Martinez

HIGHLIGHT: In this work, in order to generalize backpropagating refinement for a wide range of dense prediction tasks, we introduce a set of G-BRS (Generalized Backpropagating Refinement Scheme) layers that enable both global and localized refinement for the following tasks: interactive segmentation, semantic segmentation, image matting and monocular depth estimation.

1221, TITLE: Unsupervised Pre-Training for Temporal Action Localization Tasks

http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Unsupervised_Pre-Training_for_Temporal_Action_Localization_Tasks_CVPR_2022_paper.html

AUTHORS: Can Zhang, Tianyu Yang, Junwu Weng, Meng Cao, Jue Wang, Yuexian Zou

HIGHLIGHT: These pre-trained models can be sub-optimal for temporal localization tasks due to the inherent discrepancy between video-level classification and clip-level localization. To bridge this gap, we make the first attempt to propose a self-supervised pretext task, coined as Pseudo Action Localization (PAL) to Unsupervisedly Pre-train feature encoders for Temporal Action Localization tasks (UP-TAL).

1222, TITLE: Light Field Neural Rendering

http://openaccess.thecvf.com//content/CVPR2022/html/Suhail_Light_Field_Neural_Rendering_CVPR_2022_paper.html

AUTHORS: Mohammed Suhail, Carlos Esteves, Leonid Sigal, Ameesh Makadia

HIGHLIGHT: Methods based on geometric reconstruction need only sparse views, but cannot accurately model non-Lambertian effects. We introduce a model that combines the strengths and mitigates the limitations of these two directions.

1223, TITLE: Fast Point Transformer

http://openaccess.thecvf.com//content/CVPR2022/html/Park_Fast_Point_Transformer_CVPR_2022_paper.html

AUTHORS: Chunghyun Park, Yoonwoo Jeong, Minsu Cho, Jaesik Park

HIGHLIGHT: This paper introduces Fast Point Transformer that consists of a new lightweight self-attention layer.

1224, TITLE: Look Outside the Room: Synthesizing a Consistent Long-Term 3D Scene Video From a Single Image

http://openaccess.thecvf.com//content/CVPR2022/html/Ren_Look_Outside_the_Room_Synthesizing_a_Consistent_Long-Term_3D_Scene_CVPR_2022_paper.html

AUTHORS: Xuanchi Ren, Xiaolong Wang

HIGHLIGHT: In this paper, we propose a novel approach to synthesize a consistent long-term video given a single scene image and a trajectory of large camera motions.

1225, TITLE: Unimodal-Concentrated Loss: Fully Adaptive Label Distribution Learning for Ordinal Regression

http://openaccess.thecvf.com//content/CVPR2022/html/Li_Unimodal-Concentrated_Loss_Fully_Adaptive_Label_Distribution_Learning_for_Ordinal_Regression_CVPR_2022_paper.html

AUTHORS: Qiang Li, Jingjing Wang, Zhaoliang Yao, Yachun Li, Pengju Yang, Jingwei Yan, Chunmao Wang, Shiliang Pu

HIGHLIGHT: In this paper, we emphatically summarize that learning an adaptive label distribution on ordinal regression tasks should follow three principles.

1226, TITLE: Augmented Geometric Distillation for Data-Free Incremental Person ReID

http://openaccess.thecvf.com//content/CVPR2022/html/Lu_Augmented_Geometric_Distillation_for_Data-Free_Incremental_Person_ReID_CVPR_2022_paper.html

AUTHORS: Yichen Lu, Mei Wang, Weihong Deng

HIGHLIGHT: However, due to the strict privacy licenses and the open-set retrieval setting, it is intractable to adapt existing class IL methods to ReID. In this work, we propose an Augmented Geometric Distillation (AGD) framework to tackle these issues.

1227, TITLE: Deep Stereo Image Compression via Bi-Directional Coding

http://openaccess.thecvf.com//content/CVPR2022/html/Lei_Deep_Stereo_Image_Compression_via_Bi-Directional_Coding_CVPR_2022_paper.html

AUTHORS: Jianjun Lei, Xiangrui Liu, Bo Peng, Dengchao Jin, Wanqing Li, Jingxiao Gu

HIGHLIGHT: This paper proposes a novel bi-directional coding-based end-to-end stereo image compression network (BCSIC-Net).

1228, TITLE: Come-Closer-Diffuse-Faster: Accelerating Conditional Diffusion Models for Inverse Problems Through Stochastic Contraction

http://openaccess.thecvf.com//content/CVPR2022/html/Chung_Come-Closer-Diffuse-Faster_Accelerating_Conditional_Diffusion_Models_for_Inverse_Problems_Through_Stochastic_CVPR_2022_paper.html

AUTHORS: Hyungjin Chung, Byeongsu Sim, Jong Chul Ye

HIGHLIGHT: In this work, we show that starting from Gaussian noise is unnecessary.

1229, TITLE:	Smooth-Swap: A Simple Enhancement for Face-Swapping With Smoothness
http://openaccess.thecvf.com//content/CVPR2022/html/Kim_Smooth-Swap_A_Simple_Enhancement_for_Face-Swapping_With_Smoothness_CVPR_2022_paper.html
AUTHORS:	Jiseob Kim, Jihoon Lee, Byoung-Tak Zhang
HIGHLIGHT:	We propose a new face-swapping model called 'Smooth-Swap', which excludes complex handcrafted designs and allows fast and stable training.

1230, TITLE:	Full-Range Virtual Try-On With Recurrent Tri-Level Transform
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Full-Range_Virtual_Try-On_With_Recurrent_Tri-Level_Transform_CVPR_2022_paper.html
AUTHORS:	Han Yang, Xinrui Yu, Ziwei Liu
HIGHLIGHT:	In this work, we propose a principled framework, Recurrent Tri-Level Transform (RT-VTON), that performs full-range virtual try-on on both standard and non-standard clothes.

1231, TITLE:	Style Neophile: Constantly Seeking Novel Styles for Domain Generalization
http://openaccess.thecvf.com//content/CVPR2022/html/Kang_Style_Neophile_Constantly_Seeking_Novel_Styles_for_Domain_Generalization_CVPR_2022_paper.html
AUTHORS:	Juwon Kang, Sohyun Lee, Namyup Kim, Suha Kwak
HIGHLIGHT:	However, these methods are restricted to a finite set of styles since they obtain styles for augmentation from a fixed set of external images or by interpolating those of training data. To address this limitation and maximize the benefit of style augmentation, we propose a new method that synthesizes novel styles constantly during training.

1232, TITLE:	High-Fidelity Human Avatars From a Single RGB Camera
http://openaccess.thecvf.com//content/CVPR2022/html/Zhao_High-Fidelity_Human_Avatars_From_a_Single_RGB_Camera_CVPR_2022_paper.html
AUTHORS:	Hao Zhao, Jinsong Zhang, Yu-Kun Lai, Zerong Zheng, Yingdi Xie, Yebin Liu, Kun Li
HIGHLIGHT:	In this paper, we propose a coarse-to-fine framework to reconstruct a personalized high-fidelity human avatar from a monocular video.

1233, TITLE:	ADAPT: Vision-Language Navigation With Modality-Aligned Action Prompts
http://openaccess.thecvf.com//content/CVPR2022/html/Lin_ADAPT_Vision-Language_Navigation_With_Modality-Aligned_Action_Prompts_CVPR_2022_paper.html
AUTHORS:	Bingqian Lin, Yi Zhu, Zicong Chen, Xiwen Liang, Jianzhuang Liu, Xiaodan Liang
HIGHLIGHT:	In this paper, we propose modAlity-aligneD Action PrompTs (ADAPT), which provides the VLN agent with action prompts to enable the explicit learning of action-level modality alignment to pursue successful navigation.

1234, TITLE:	Multiview Transformers for Video Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Yan_Multiview_Transformers_for_Video_Recognition_CVPR_2022_paper.html
AUTHORS:	Shen Yan, Xuehan Xiong, Anurag Arnab, Zhichao Lu, Mi Zhang, Chen Sun, Cordelia Schmid
HIGHLIGHT:	Although transformer architectures have recently advanced the state-of-the-art, they have not explicitly modelled different spatiotemporal resolutions. To this end, we present Multiview Transformers for Video Recognition (MTV).

1235, TITLE:	RIO: Rotation-Equivariance Supervised Learning of Robust Inertial Odometry
http://openaccess.thecvf.com//content/CVPR2022/html/Cao_RIO_Rotation-Equivariance_Supervised_Learning_of_Robust_Inertial_Odometry_CVPR_2022_paper.html
AUTHORS:	Xiya Cao, Caifa Zhou, Dandan Zeng, Yongliang Wang
HIGHLIGHT:	This paper introduces rotation-equivariance as a self-supervisor to train inertial odometry models.

1236, TITLE:	How Good Is Aesthetic Ability of a Fashion Model?
http://openaccess.thecvf.com//content/CVPR2022/html/Zou_How_Good_Is_Aesthetic_Ability_of_a_Fashion_Model_CVPR_2022_paper.html
AUTHORS:	Xingxing Zou, Kaicheng Pang, Wen Zhang, Waikeung Wong
HIGHLIGHT:	We introduce A100 (Aesthetic 100) to assess the aesthetic ability of the fashion compatibility models.

1237, TITLE:	Mining Multi-View Information: A Strong Self-Supervised Framework for Depth-Based 3D Hand Pose and Mesh Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Ren_Mining_Multi-View_Information_A_Strong_Self-Supervised_Framework_for_Depth-Based_3D_CVPR_2022_paper.html
AUTHORS:	Pengfei Ren, Haifeng Sun, Jiachang Hao, Jingyu Wang, Qi Qi, Jianxin Liao
HIGHLIGHT:	In this work, we study the cross-view information fusion problem in the task of self-supervised 3D hand pose estimation from the depth image.

1238, TITLE:	Automated Progressive Learning for Efficient Training of Vision Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Automated_Progressive_Learning_for_Efficient_Training_of_Vision_Transformers_CVPR_2022_paper.html
AUTHORS:	Changlin Li, Bohan Zhuang, Guangrun Wang, Xiaodan Liang, Xiaojun Chang, Yi Yang
HIGHLIGHT:	In this paper, we take a practical step towards efficient training of ViTs by customizing and automating progressive learning.

1239, TITLE: BTS: A Bi-Lingual Benchmark for Text Segmentation in the Wild
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_BTS_A_Bi-Lingual_Benchmark_for_Text_Segmentation_in_the_Wild_CVPR_2022_paper.html
AUTHORS: Xixi Xu, Zhongang Qi, Jianqi Ma, Honglun Zhang, Ying Shan, Xiaohu Qie
HIGHLIGHT: Different from English which has a limited alphabet of letters, Chinese has much more basic characters with complex structures, making the problem more difficult to deal with. To better analyze this problem, we propose the Bi-lingual Text Segmentation (BTS) dataset, a benchmark that covers various common Chinese scenes including 14,250 diverse and fine-annotated text images.

1240, TITLE: Learning Structured Gaussians To Approximate Deep Ensembles
http://openaccess.thecvf.com//content/CVPR2022/html/Simpson_Learning_Structured_Gaussians_To_Approximate_Deep_Ensembles_CVPR_2022_paper.html
AUTHORS: Ivor J. A. Simpson, Sara Vicente, Neill D. F. Campbell
HIGHLIGHT: This paper proposes using a sparse-structured multivariate Gaussian to provide a closed-form approximator for the output of probabilistic ensemble models used for dense image prediction tasks.

1241, TITLE: Adaptive Trajectory Prediction via Transferable GNN
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_Adaptive_Trajectory_Prediction_via_Transferable_GNN_CVPR_2022_paper.html
AUTHORS: Yi Xu, Lichen Wang, Yizhou Wang, Yun Fu
HIGHLIGHT: This issue results in inevitable performance decrease. To address this issue, we propose a novel Transferable Graph Neural Network (T-GNN) framework, which jointly conducts trajectory prediction as well as domain alignment in a unified framework.

1242, TITLE: Total Variation Optimization Layers for Computer Vision
http://openaccess.thecvf.com//content/CVPR2022/html/Yeh_Total_Variation_Optimization_Layers_for_Computer_Vision_CVPR_2022_paper.html
AUTHORS: Raymond A. Yeh, Yuan-Ting Hu, Zhongzheng Ren, Alexander G. Schwing
HIGHLIGHT: To study question (a), in this work, we propose total variation (TV) minimization as a layer for computer vision.

1243, TITLE: Defensive Patches for Robust Recognition in the Physical World
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Defensive_Patches_for_Robust_Recognition_in_the_Physical_World_CVPR_2022_paper.html
AUTHORS: Jiakai Wang, Zixin Yin, Pengfei Hu, Aishan Liu, Renshuai Tao, Haotong Qin, Xianglong Liu, Dacheng Tao
HIGHLIGHT: Motivated by the fact that robust recognition depends on both local and global features, we propose a defensive patch generation framework to address these problems by helping models better exploit these features.

1244, TITLE: Single-Stage Is Enough: Multi-Person Absolute 3D Pose Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Jin_Single-Stage_Is_Enough_Multi-Person_Absolute_3D_Pose_Estimation_CVPR_2022_paper.html
AUTHORS: Lei Jin, Chenyang Xu, Xiaojuan Wang, Yabo Xiao, Yandong Guo, Xuecheng Nie, Jian Zhao
HIGHLIGHT: We argue that it is more desirable to simplify such two-stage paradigm to a single-stage one to promote both efficiency and performance. To this end, we present an efficient single-stage solution, Decoupled Regression Model (DRM), with three distinct novelties.

1245, TITLE: Deformation and Correspondence Aware Unsupervised Synthetic-to-Real Scene Flow Estimation for Point Clouds
http://openaccess.thecvf.com//content/CVPR2022/html/Jin_Deformation_and_Correspondence_Aware_Unsupervised_Synthetic-to-Real_Scene_Flow_Estimation_for_CVPR_2022_paper.html
AUTHORS: Zhao Jin, Yinjie Lei, Naveed Akhtar, Haifeng Li, Munawar Hayat
HIGHLIGHT: However, large disparities between existing synthetic datasets and real scenes lead to poor model transfer. We make two major contributions to address that.

1246, TITLE: Learn From Others and Be Yourself in Heterogeneous Federated Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Huang_Learn_From_Others_and_Be_Yourself_in_Heterogeneous_Federated_Learning_CVPR_2022_paper.html
AUTHORS: Wenke Huang, Mang Ye, Bo Du
HIGHLIGHT: In this work, we propose FCCL (Federated Cross-Correlation and Continual Learning).

1247, TITLE: Sequential Voting With Relational Box Fields for Active Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Fu_Sequential_Voting_With_Relational_Box_Fields_for_Active_Object_Detection_CVPR_2022_paper.html
AUTHORS: Qichen Fu, Xingyu Liu, Kris Kitani
HIGHLIGHT: To leverage each pixel as evidence to determine the bounding box of the active object, we propose a pixel-wise voting function.

1248, TITLE: Semantic-Aware Auto-Encoders for Self-Supervised Representation Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Semantic-Aware_Auto-Encoders_for_Self-Supervised_Representation_Learning_CVPR_2022_paper.html

AUTHORS: Guangrun Wang, Yansong Tang, Liang Lin, Philip H.S. Torr
HIGHLIGHT: In this paper, we switch from D models to G models using the classical auto-encoder (AE).

1249, TITLE: Learning Transferable Human-Object Interaction Detector With Natural Language Supervision
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Learning_Transferable_Human-
Object_Interaction_Detector_With_Natural_Language_Supervision_CVPR_2022_paper.html
AUTHORS: Suchen Wang, Yueqi Duan, Henghui Ding, Yap-Peng Tan, Kim-Hui Yap, Junsong Yuan
HIGHLIGHT: In this work, we aim to develop a transferable HOI detector for unseen interactions.

1250, TITLE: Fourier Document Restoration for Robust Document Dewarping and Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Xue_Fourier_Document_Restoration_for_Robust_Document_Dewarping_and
_Recognition_CVPR_2022_paper.html
AUTHORS: Chuhui Xue, Zichen Tian, Fangneng Zhan, Shijian Lu, Song Bai
HIGHLIGHT: This paper presents FDRNet, a Fourier Document Restoration Network that can restore documents with
different distortions and improve document recognition in a reliable and simpler manner.

1251, TITLE: Consistency Learning via Decoding Path Augmentation for Transformers in Human Object Interaction
Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Park_Consistency_Learning_via_Decoding_Path_Augmentation_for_Transfor
mers_in_Human_CVPR_2022_paper.html
AUTHORS: Jihwan Park, SeungJun Lee, Hwan Heo, Hyeong Kyu Choi, Hyunwoo J. Kim
HIGHLIGHT: Motivated by various inference paths for HOI detection, we propose cross-path consistency learning (CPC),
which is a novel end-to-end learning strategy to improve HOI detection for transformers by leveraging augmented decoding paths.

1252, TITLE: Consistent Explanations by Contrastive Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Pillai_Consistent_Explanations_by_Contrastive_Learning_CVPR_2022_paper
.html
AUTHORS: Vipin Pillai, Soroush Abbasi Koohpayegani, Ashley Ouligian, Dennis Fong, Hamed Pirsiavash
HIGHLIGHT: Given an interpretation algorithm, e.g., Grad-CAM, we introduce a novel training method to train the model to
produce more consistent explanations.

1253, TITLE: Text2Pos: Text-to-Point-Cloud Cross-Modal Localization
http://openaccess.thecvf.com//content/CVPR2022/html/Kolmet_Text2Pos_Text-to-Point-Cloud_Cross-
Modal_Localization_CVPR_2022_paper.html
AUTHORS: Manuel Kolmet, Qunjie Zhou, Aljoša Ošep, Laura Leal-Taixé
HIGHLIGHT: In particular, we propose Text2Pos, a cross-modal localization module that learns to align textual descriptions
with localization cues in a coarse- to-fine manner.

1254, TITLE: MulT: An End-to-End Multitask Learning Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Bhattacharjee_MulT_An_End-to-
End_Multitask_Learning_Transformer_CVPR_2022_paper.html
AUTHORS: Deblina Bhattacharjee, Tong Zhang, Sabine Süsstrunk, Mathieu Salzmann
HIGHLIGHT: We propose an end-to-end Multitask Learning Transformer framework, named MulT, to simultaneously learn
multiple high-level vision tasks, including depth estimation, semantic segmentation, reshading, surface normal estimation, 2D
keypoint detection, and edge detection.

1255, TITLE: Hierarchical Modular Network for Video Captioning
http://openaccess.thecvf.com//content/CVPR2022/html/Ye_Hierarchical_Modular_Network_for_Video_Captioning_CVPR_2022_pa
per.html
AUTHORS: Hanhua Ye, Guorong Li, Yuankai Qi, Shuhui Wang, Qingming Huang, Ming-Hsuan Yang
HIGHLIGHT: In this work, we propose a hierarchical modular network to bridge video representations and linguistic
semantics from three levels before generating captions.

1256, TITLE: Learning With Neighbor Consistency for Noisy Labels
http://openaccess.thecvf.com//content/CVPR2022/html/Iscen_Learning_With_Neighbor_Consistency_for_Noisy_Labels_CVPR_202
2_paper.html
AUTHORS: Ahmet Iscen, Jack Valmadre, Anurag Arnab, Cordelia Schmid
HIGHLIGHT: However, collecting large datasets in a time- and cost-efficient manner often results in label noise. We present a
method for learning from noisy labels that leverages similarities between training examples in feature space, encouraging the
prediction of each example to be similar to its nearest neighbours.

1257, TITLE: Depth Estimation by Combining Binocular Stereo and Monocular Structured-Light
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_Depth_Estimation_by_Combining_Binocular_Stereo_and_Monocular_Str
uctured-Light_CVPR_2022_paper.html
AUTHORS: Yuhua Xu, Xiaoli Yang, Yushan Yu, Wei Jia, Zhaobi Chu, Yulan Guo
HIGHLIGHT: In this paper, we present a novel stereo system, which consists of two cameras (an RGB camera and an IR
camera) and an IR speckle projector.

1258, TITLE: Salient-to-Broad Transition for Video Person Re-Identification

http://openaccess.thecvf.com//content/CVPR2022/html/Bai_Salient-to-Broad_Transition_for_Video_Person_Re-Identification_CVPR_2022_paper.html
AUTHORS: Shutao Bai, Bingpeng Ma, Hong Chang, Rui Huang, Xilin Chen
HIGHLIGHT: Due to the limited utilization of temporal relations in video re-id, the frame-level attention regions of mainstream methods are partial and highly similar. To address this problem, we propose a Salient-to-Broad Module (SBM) to enlarge the attention regions gradually.

1259, TITLE: Object-Region Video Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Herzig_Object-Region_Video_Transformers_CVPR_2022_paper.html
AUTHORS: Roei Herzig, Elad Ben-Avraham, Karttikeya Mangalam, Amir Bar, Gal Chechik, Anna Rohrbach, Trevor Darrell, Amir Globerson
HIGHLIGHT: In this work, we present Object-Region Video Transformers (ORViT), an object-centric approach that extends video transformer layers with a block that directly incorporates object representations.

1260, TITLE: DeeCap: Dynamic Early Exiting for Efficient Image Captioning
http://openaccess.thecvf.com//content/CVPR2022/html/Fei_DeeCap_Dynamic_Early_Exiting_for_Efficient_Image_Captioning_CVPR_2022_paper.html
AUTHORS: Zhengcong Fei, Xu Yan, Shuhui Wang, Qi Tian
HIGHLIGHT: On the other hand, the exiting decisions made by internal classifiers are unreliable sometimes. To solve these issues, we propose DeeCap framework for efficient image captioning, which dynamically selects proper-sized decoding layers from a global perspective to exit early.

1261, TITLE: AME: Attention and Memory Enhancement in Hyper-Parameter Optimization
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_AME_Attention_and_Memory_Enhancement_in_Hyper-Parameter_Optimization_CVPR_2022_paper.html
AUTHORS: Nuo Xu, Jianlong Chang, Xing Nie, Chunlei Huo, Shiming Xiang, Chunhong Pan
HIGHLIGHT: Training Deep Neural Networks (DNNs) is inherently subject to sensitive hyper-parameters and untimely feedbacks of performance evaluation. To solve these two difficulties, an efficient parallel hyper-parameter optimization model is proposed under the framework of Deep Reinforcement Learning (DRL).

1262, TITLE: Alignment-Uniformity Aware Representation Learning for Zero-Shot Video Classification
http://openaccess.thecvf.com//content/CVPR2022/html/Pu_Alignment-Uniformity_Aware_Representation_Learning_for_Zero-Shot_Video_Classification_CVPR_2022_paper.html
AUTHORS: Shi Pu, Kaili Zhao, Mao Zheng
HIGHLIGHT: To enhance model generalizability, this paper presents an end-to-end framework that preserves alignment and uniformity properties for representations on both seen and unseen classes.

1263, TITLE: RepMLPNet: Hierarchical Vision MLP With Re-Parameterized Locality
http://openaccess.thecvf.com//content/CVPR2022/html/Ding_RepMLPNet_Hierarchical_Vision_MLP_With_Re-Parameterized_Locality_CVPR_2022_paper.html
AUTHORS: Xiaohan Ding, Honghao Chen, Xiangyu Zhang, Jungong Han, Guiguang Ding
HIGHLIGHT: In this paper, we propose a methodology, Locality Injection, to incorporate local priors into an FC layer via merging the trained parameters of a parallel conv kernel into the FC kernel.

1264, TITLE: DR.VIC: Decomposition and Reasoning for Video Individual Counting
http://openaccess.thecvf.com//content/CVPR2022/html/Han_DR.VIC_Decomposition_and_Reasoning_for_Video_Individual_Counting_CVPR_2022_paper.html
AUTHORS: Tao Han, Lei Bai, Junyu Gao, Qi Wang, Wanli Ouyang
HIGHLIGHT: In this work, we propose to conduct the pedestrian counting from a new perspective - Video Individual Counting (VIC), which counts the total number of individual pedestrians in the given video (a person is only counted once).

1265, TITLE: LiDARCap: Long-Range Marker-Less 3D Human Motion Capture With LiDAR Point Clouds
http://openaccess.thecvf.com//content/CVPR2022/html/Li_LiDARCap_Long-Range_Marker-Less_3D_Human_Motion_Capture_With_LiDAR_Point_CVPR_2022_paper.html
AUTHORS: Jialian Li, Jingyi Zhang, Zhiyong Wang, Siqi Shen, Chenglu Wen, Yuexin Ma, Lan Xu, Jingyi Yu, Cheng Wang
HIGHLIGHT: Existing motion capture datasets are largely short-range and cannot yet fit the need of long-range applications. We propose LiDARHuman26M, a new human motion capture dataset captured by LiDAR at a much longer range to overcome this limitation.

1266, TITLE: GeoEngine: A Platform for Production-Ready Geospatial Research
http://openaccess.thecvf.com//content/CVPR2022/html/Verma_GeoEngine_A_Platform_for_Production-Ready_Geospatial_Research_CVPR_2022_paper.html
AUTHORS: Sagar Verma, Siddharth Gupta, Hal Shin, Akash Panigrahi, Shubham Goswami, Shweta Pardeshi, Natanael Exe, Ujwal Dutta, Tanka Raj Joshi, Nitin Bhojwani
HIGHLIGHT: In this paper we introduce the GeoEngine platform for reproducible and production-ready geospatial machine learning research.

1267, TITLE: Revisiting Document Image Dewarping by Grid Regularization

http://openaccess.thecvf.com//content/CVPR2022/html/Jiang_Revisiting_Document_Image_Dewarping_by_Grid_Regularization_CVPR_2022_paper.html
AUTHORS: Xiangwei Jiang, Rujiao Long, Nan Xue, Zhibo Yang, Cong Yao, Gui-Song Xia
HIGHLIGHT: This paper addresses the problem of document image dewarping, which aims at eliminating the geometric distortion in document images for document digitization.

1268, TITLE: Semi-Supervised Few-Shot Learning via Multi-Factor Clustering
http://openaccess.thecvf.com//content/CVPR2022/html/Ling_Semi-Supervised_Few-Shot_Learning_via_Multi-Factor_Clustering_CVPR_2022_paper.html
AUTHORS: Jie Ling, Lei Liao, Meng Yang, Jia Shuai
HIGHLIGHT: However, the relationship between unlabeled and labeled data is not well exploited in generating pseudo labels, the noise of which will directly harm the model learning. In this paper, we propose a Clustering-based semi-supervised Few-Shot Learning (cluster-FSL) method to solve the above problems in image classification.

1269, TITLE: CMT-DeepLab: Clustering Mask Transformers for Panoptic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Yu_CMT-DeepLab_Clustering_Mask_Transformers_for_Panoptic_Segmentation_CVPR_2022_paper.html
AUTHORS: Qihang Yu, Huiyu Wang, Dahun Kim, Siyuan Qiao, Maxwell Collins, Yukun Zhu, Hartwig Adam, Alan Yuille, Liang-Chieh Chen
HIGHLIGHT: We propose Clustering Mask Transformer (CMT-DeepLab), a transformer-based framework for panoptic segmentation designed around clustering.

1270, TITLE: Weakly-Supervised Generation and Grounding of Visual Descriptions With Conditional Generative Models
http://openaccess.thecvf.com//content/CVPR2022/html/Mavroudi_Weakly-Supervised_Generation_and_Grounding_of_Visual_Descriptions_With_Conditional_Generative_CVPR_2022_paper.html
AUTHORS: Effrosyni Mavroudi, René Vidal
HIGHLIGHT: However, this leads to sub-optimal grounding, since attention coefficients are computed without taking into account the word that needs to be localized. To address this shortcoming, we propose a novel Grounded Visual Description Conditional Variational Autoencoder (GVD-CVAE) and leverage its latent variables for grounding.

1271, TITLE: Novel Class Discovery in Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Zhao_Novel_Class_Discovery_in_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Yuyang Zhao, Zhun Zhong, Nicu Sebe, Gim Hee Lee
HIGHLIGHT: We introduce a new setting of Novel Class Discovery in Semantic Segmentation (NCDSS), which aims at segmenting unlabeled images containing new classes given prior knowledge from a labeled set of disjoint classes.

1272, TITLE: ARCS: Accurate Rotation and Correspondence Search
http://openaccess.thecvf.com//content/CVPR2022/html/Peng_ARCS_Accurate_Rotation_and_Correspondence_Search_CVPR_2022_paper.html
AUTHORS: Liangzu Peng, Manolis C. Tsakiris, René Vidal
HIGHLIGHT: This paper is about the old Wahba problem in its more general form, which we call "simultaneous rotation and correspondence search".

1273, TITLE: Learning To Anticipate Future With Dynamic Context Removal
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_Learning_To_Anticipate_Future_With_Dynamic_Context_Removal_CVPR_2022_paper.html
AUTHORS: Xinyu Xu, Yong-Lu Li, Cewu Lu
HIGHLIGHT: In this filed, previous methods usually care more about the model architecture design or but few attention has been put on how to train an anticipation model with a proper learning policy. To this end, in this work, we propose a novel training scheme called Dynamic Context Removal (DCR), which dynamically schedule the visibility of observed future in the learning procedure.

1274, TITLE: GCFSR: A Generative and Controllable Face Super Resolution Method Without Facial and GAN Priors
http://openaccess.thecvf.com//content/CVPR2022/html/He_GCFSR_A_Generative_and_Controllable_Face_Super_Resolution_Method_Without_CVPR_2022_paper.html
AUTHORS: Jingwen He, Wu Shi, Kai Chen, Lean Fu, Chao Dong
HIGHLIGHT: In this work, we propose a generative and controllable face SR framework, called GCFSR, which can reconstruct images with faithful identity information without any additional priors.

1275, TITLE: Perception Prioritized Training of Diffusion Models
http://openaccess.thecvf.com//content/CVPR2022/html/Choi_Perception_Prioritized_Training_of_Diffusion_Models_CVPR_2022_paper.html
AUTHORS: Jooyoung Choi, Jungbeom Lee, Chaehun Shin, Sungwon Kim, Hyunwoo Kim, Sungroh Yoon
HIGHLIGHT: In this paper, we show that restoring data corrupted with certain noise levels offers a proper pretext task for the model to learn rich visual concepts.

1276, TITLE: Using 3D Topological Connectivity for Ghost Particle Reduction in Flow Reconstruction
http://openaccess.thecvf.com//content/CVPR2022/html/Tsalicoglou_Using_3D_Topological_Connectivity_for_Ghost_Particle_Reduction_in_Flow_CVPR_2022_paper.html

AUTHORS: Christina Tsalicoglou, Thomas Rösgen
HIGHLIGHT: In an alternative method, 3D Particle Streak Velocimetry (3D-PSV), the exposure time is increased, and the particles' pathlines are imaged as "streaks". We treat these streaks (a) as connected endpoints and (b) as conic section segments and develop a theoretical model that describes the mechanisms of 3D ambiguity generation and shows that streaks can drastically reduce reconstruction ambiguities.

1277, TITLE: On the Integration of Self-Attention and Convolution
http://openaccess.thecvf.com//content/CVPR2022/html/Pan_On_the_Integration_of_Self-Attention_and_Convolution_CVPR_2022_paper.html
AUTHORS: Xuran Pan, Chunjiang Ge, Rui Lu, Shiji Song, Guanfu Chen, Zeyi Huang, Gao Huang
HIGHLIGHT: Convolution and self-attention are two powerful techniques for representation learning, and they are usually considered as two peer approaches that are distinct from each other. In this paper, we show that there exists a strong underlying relation between them, in the sense that the bulk of computations of these two paradigms are in fact done with the same operation.

1278, TITLE: Progressively Generating Better Initial Guesses Towards Next Stages for High-Quality Human Motion Prediction
http://openaccess.thecvf.com//content/CVPR2022/html/Ma_Progressively_Generating_Better_Initial_Guesses_Towards_Next_Stages_for_High-Quality_CVPR_2022_paper.html
AUTHORS: Tiezheng Ma, Yongwei Nie, Chengjiang Long, Qing Zhang, Guiqing Li
HIGHLIGHT: This paper presents a high-quality human motion prediction method that accurately predicts future human poses given observed ones.

1279, TITLE: CHEX: CHannel EXploration for CNN Model Compression
http://openaccess.thecvf.com//content/CVPR2022/html/Hou_CHEX_CHannel_EXploration_for_CNN_Model_Compression_CVPR_2022_paper.html
AUTHORS: Zejiang Hou, Minghai Qin, Fei Sun, Xiaolong Ma, Kun Yuan, Yi Xu, Yen-Kuang Chen, Rong Jin, Yuan Xie, Sun-Yuan Kung
HIGHLIGHT: Such limitations may lead to sub-optimal model quality as well as excessive memory and training cost. In this paper, we propose a novel Channel Exploration methodology, dubbed as CHEX, to rectify these problems.

1280, TITLE: M2I: From Factored Marginal Trajectory Prediction to Interactive Prediction
http://openaccess.thecvf.com//content/CVPR2022/html/Sun_M2I_From_Factored_Marginal_Trajectory_Prediction_to_Interactive_Prediction_CVPR_2022_paper.html
AUTHORS: Qiao Sun, Xin Huang, Junru Gu, Brian C. Williams, Hang Zhao
HIGHLIGHT: In this work, we exploit the underlying relations between interacting agents and decouple the joint prediction problem into marginal prediction problems.

1281, TITLE: Domain Adaptation on Point Clouds via Geometry-Aware Implicits
http://openaccess.thecvf.com//content/CVPR2022/html/Shen_Domain_Adaptation_on_Point_Clouds_via_Geometry-Aware_Implicits_CVPR_2022_paper.html
AUTHORS: Yuefan Shen, Yanchao Yang, Mi Yan, He Wang, Youyi Zheng, Leonidas J. Guibas
HIGHLIGHT: Here we propose a simple yet effective method for unsupervised domain adaptation on point clouds by employing a self-supervised task of learning geometry-aware implicits, which plays two critical roles in one shot.

1282, TITLE: Consistency Driven Sequential Transformers Attention Model for Partially Observable Scenes
http://openaccess.thecvf.com//content/CVPR2022/html/Rangrej_Consistency_Driven_Sequential_Transformers_Attention_Model_for_Partially_Observable_Scenes_CVPR_2022_paper.html
AUTHORS: Samrudhdhi B. Rangrej, Chetan L. Srinidhi, James J. Clark
HIGHLIGHT: In this paper, we develop a Sequential Transformers Attention Model (STAM) that only partially observes a complete image and predicts informative glimpse locations solely based on past glimpses.

1283, TITLE: GroupViT: Semantic Segmentation Emerges From Text Supervision
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_GroupViT_Semantic_Segmentation_Emerges_From_Text_Supervision_CVPR_2022_paper.html
AUTHORS: Jiarui Xu, Shalini De Mello, Sifei Liu, Wonmin Byeon, Thomas Breuel, Jan Kautz, Xiaolong Wang
HIGHLIGHT: Instead, in this paper, we propose to bring back the grouping mechanism into deep networks, which allows semantic segments to emerge automatically with only text supervision.

1284, TITLE: NeuralHOFusion: Neural Volumetric Rendering Under Human-Object Interactions
http://openaccess.thecvf.com//content/CVPR2022/html/Jiang_NeuralHOFusion_Neural_Volumetric_Rendering_Under_Human-Object_Interactions_CVPR_2022_paper.html
AUTHORS: Yuheng Jiang, Suyi Jiang, Guoxing Sun, Zhuo Su, Kaiwen Guo, Minye Wu, Jingyi Yu, Lan Xu
HIGHLIGHT: In this paper, we propose NeuralHOFusion, a neural approach for volumetric human-object capture and rendering using sparse consumer RGBD sensors.

1285, TITLE: Generalizable Human Pose Triangulation
http://openaccess.thecvf.com//content/CVPR2022/html/Bartol_Generalizable_Human_Pose_Triangulation_CVPR_2022_paper.html
AUTHORS: Kristijan Bartol, David BojaniÄ‡, Tomislav PetkoviÄ‡, Tomislav PribaniÄ‡
HIGHLIGHT: We propose a stochastic framework for human pose triangulation and demonstrate a superior generalization across different camera arrangements on two public datasets.

1286, TITLE: DiffusionCLIP: Text-Guided Diffusion Models for Robust Image Manipulation
http://openaccess.thecvf.com//content/CVPR2022/html/Kim_DiffusionCLIP_Text-
Guided_Diffusion_Models_for_Robust_Image_Manipulation_CVPR_2022_paper.html
AUTHORS: Gwanghyun Kim, Taesung Kwon, Jong Chul Ye
HIGHLIGHT: Specifically, these approaches often have difficulties in reconstructing images with novel poses, views, and highly variable contents compared to the training data, altering object identity, or producing unwanted image artifacts. To mitigate these problems and enable faithful manipulation of real images, we propose a novel method, dubbed DiffusionCLIP, that performs text-driven image manipulation using diffusion models.

1287, TITLE: Occlusion-Aware Cost Constructor for Light Field Depth Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Occlusion-
Aware_Cost_Constructor_for_Light_Field_Depth_Estimation_CVPR_2022_paper.html
AUTHORS: Yingqian Wang, Longguang Wang, Zhengyu Liang, Jungang Yang, Wei An, Yulan Guo
HIGHLIGHT: In this paper, we propose a simple and fast cost constructor to construct matching cost for LF depth estimation.

1288, TITLE: SmartPortraits: Depth Powered Handheld Smartphone Dataset of Human Portraits for State Estimation, Reconstruction and Synthesis
http://openaccess.thecvf.com//content/CVPR2022/html/Kornilova_SmartPortraits_Depth_Powered_Handheld_Smartphone_Dataset_of_Human_Portraits_for_CVPR_2022_paper.html
AUTHORS: Anastasiia Kornilova, Marsel Faizullin, Konstantin Pakulev, Andrey Sadkov, Denis Kukushkin, Azat Akhmetyanov, Timur Akhtyamov, Hekmat Taherinejad, Gonzalo Ferrer
HIGHLIGHT: We present a dataset of 1000 video sequences of human portraits recorded in real and uncontrolled conditions by using a handheld smartphone accompanied by an external high-quality depth camera.

1289, TITLE: BppAttack: Stealthy and Efficient Trojan Attacks Against Deep Neural Networks via Image Quantization and Contrastive Adversarial Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_BppAttack_Stealthy_and_Efficient_Trojan_Attacks_Against_Deep_Neural_Networks_CVPR_2022_paper.html
AUTHORS: Zhenting Wang, Juan Zhai, Shiqing Ma
HIGHLIGHT: In this paper, we propose stealthy and efficient Trojan attacks, BppAttack.

1290, TITLE: GlideNet: Global, Local and Intrinsic Based Dense Embedding NETwork for Multi-Category Attributes Prediction
http://openaccess.thecvf.com//content/CVPR2022/html/Metwaly_GlideNet_Global_Local_and_Intrinsic_Based_Dense_Embedding_NETwork_for_CVPR_2022_paper.html
AUTHORS: Kareem Metwaly, Aerin Kim, Elliot Branson, Vishal Monga
HIGHLIGHT: Yet significant challenges remain in: 1) predicting a large number of attributes over multiple object categories, 2) modeling category-dependence of attributes, 3) methodically capturing both global and local scene context, and 4) robustly predicting attributes of objects with low pixel-count. To address these issues, we propose a novel multi-category attribute prediction deep architecture named GlideNet, which contains three distinct feature extractors.

1291, TITLE: Stacked Hybrid-Attention and Group Collaborative Learning for Unbiased Scene Graph Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Dong_Stacked_Hybrid-
Attention_and_Group_Collaborative_Learning_for_Unbiased_Scene_Graph_CVPR_2022_paper.html
AUTHORS: Xingning Dong, Tian Gan, Xuemeng Song, Jianlong Wu, Yuan Cheng, Liqiang Nie
HIGHLIGHT: Towards this end, in this paper, we first present a novel Stacked Hybrid-Attention network, which facilitates the intra-modal refinement as well as the inter-modal interaction, to serve as the encoder. We then devise an innovative Group Collaborative Learning strategy to optimize the decoder.

1292, TITLE: Ensembling Off-the-Shelf Models for GAN Training
http://openaccess.thecvf.com//content/CVPR2022/html/Kumari_Ensembling_Off-the-
Shelf_Models_for_GAN_Training_CVPR_2022_paper.html
AUTHORS: Nupur Kumari, Richard Zhang, Eli Shechtman, Jun-Yan Zhu
HIGHLIGHT: We propose an effective selection mechanism, by probing the linear separability between real and fake samples in pretrained model embeddings, choosing the most accurate model, and progressively adding it to the discriminator ensemble.

1293, TITLE: Towards Better Plasticity-Stability Trade-Off in Incremental Learning: A Simple Linear Connector
http://openaccess.thecvf.com//content/CVPR2022/html/Lin_Towards_Better_Plasticity-Stability_Trade-
Off_in_Incremental_Learning_A_Simple_Linear_CVPR_2022_paper.html
AUTHORS: Guoliang Lin, Hanlu Chu, Hanjiang Lai
HIGHLIGHT: In this work, we propose to employ mode connectivity in loss landscapes to achieve better plasticity-stability trade-off without any previous samples.

1294, TITLE: Topology-Preserving Shape Reconstruction and Registration via Neural Diffeomorphic Flow
http://openaccess.thecvf.com//content/CVPR2022/html/Sun_Topology-
Preserving_Shape_Reconstruction_and_Registration_via_Neural_Diffeomorphic_Flow_CVPR_2022_paper.html
AUTHORS: Shanlin Sun, Kun Han, Deying Kong, Hao Tang, Xiangyi Yan, Xiaohui Xie
HIGHLIGHT: In this paper, we propose a new model called Neural Diffeomorphic Flow (NDF) to learn deep implicit shape templates, representing shapes as conditional diffeomorphic deformations of templates, intrinsically preserving shape topologies.

1295, TITLE: Segment and Complete: Defending Object Detectors Against Adversarial Patch Attacks With Robust Patch Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Segment_and_Complete_Defending_Object_Detectors_Against_Adversarial_Patch_Attacks_CVPR_2022_paper.html
AUTHORS: Jiang Liu, Alexander Levine, Chun Pong Lau, Rama Chellappa, Soheil Feizi
HIGHLIGHT: In this paper, we propose Segment and Complete defense (SAC), a general framework for defending object detectors against patch attacks through detection and removal of adversarial patches.

1296, TITLE: Cross-Domain Few-Shot Learning With Task-Specific Adapters
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Cross-Domain_Few-Shot_Learning_With_Task-Specific_Adapters_CVPR_2022_paper.html
AUTHORS: Wei-Hong Li, Xialei Liu, Hakan Bilen
HIGHLIGHT: In this paper, we look at the problem of cross-domain few-shot classification that aims to learn a classifier from previously unseen classes and domains with few labeled samples.

1297, TITLE: MAXIM: Multi-Axis MLP for Image Processing
http://openaccess.thecvf.com//content/CVPR2022/html/Tu_MAXIM_Multi-Axis_MLP_for_Image_Processing_CVPR_2022_paper.html
AUTHORS: Zhengzhong Tu, Hossein Talebi, Han Zhang, Feng Yang, Peyman Milanfar, Alan Bovik, Yinxiao Li
HIGHLIGHT: In this work, we present a multi-axis MLP based architecture called MAXIM, that can serve as an efficient and flexible general-purpose vision backbone for image processing tasks.

1298, TITLE: Learning Part Segmentation Through Unsupervised Domain Adaptation From Synthetic Vehicles
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Learning_Part_Segmentation_Through_Unsupervised_Domain_Adaptation_From_Synthetic_Vehicles_CVPR_2022_paper.html
AUTHORS: Qing Liu, Adam Kortylewski, Zhishuai Zhang, Zizhang Li, Mengqi Guo, Qihao Liu, Xiaoding Yuan, Jiteng Mu, Weichao Qiu, Alan Yuille
HIGHLIGHT: In this paper, we propose the idea of learning part segmentation through unsupervised domain adaptation (UDA) from synthetic data.

1299, TITLE: Delving Into the Estimation Shift of Batch Normalization in a Network
http://openaccess.thecvf.com//content/CVPR2022/html/Huang_Delving_Into_the_Estimation_Shift_of_Batch_Normalization_in_a_CVPR_2022_paper.html
AUTHORS: Lei Huang, Yi Zhou, Tian Wang, Jie Luo, Xianglong Liu
HIGHLIGHT: This paper focuses on investigating the estimation of population statistics.

1300, TITLE: Towards Better Understanding Attribution Methods
http://openaccess.thecvf.com//content/CVPR2022/html/Rao_Towards_Better_Understanding_Attribution_Methods_CVPR_2022_paper.html
AUTHORS: Sukrut Rao, Moritz Böhle, Bernt Schiele
HIGHLIGHT: Evaluating such methods is challenging since no ground truth attributions exist. We thus propose three novel evaluation schemes to more reliably measure the faithfulness of those methods, to make comparisons between them more fair, and to make visual inspection more systematic.

1301, TITLE: Learning Object Context for Novel-View Scene Layout Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Qiao_Learning_Object_Context_for_Novel-View_Scene_Layout_Generation_CVPR_2022_paper.html
AUTHORS: Xiaotian Qiao, Gerhard P. Hancke, Rynson W.H. Lau
HIGHLIGHT: Such a problem is challenging as it involves accurate understanding of the 3D geometry and semantics of the scene from as little as a single 2D scene layout. To tackle this challenging problem, we propose a deep model to capture contextualized object representation by explicitly modeling the object context transformation in the scene.

1302, TITLE: PSTR: End-to-End One-Step Person Search With Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Cao_PSTR_End-to-End_One-Step_Person_Search_With_Transformers_CVPR_2022_paper.html
AUTHORS: Jiale Cao, Yanwei Pang, Rao Muhammad Anwer, Hisham Cholakkal, Jin Xie, Mubarak Shah, Fahad Shahbaz Khan
HIGHLIGHT: We propose a novel one-step transformer-based person search framework, PSTR, that jointly performs person detection and re-identification (re-id) in a single architecture.

1303, TITLE: Neural Fields As Learnable Kernels for 3D Reconstruction
http://openaccess.thecvf.com//content/CVPR2022/html/Williams_Neural_Fields_As_Learnable_Kernels_for_3D_Reconstruction_CVPR_2022_paper.html
AUTHORS: Francis Williams, Zan Gojcic, Sameh Khamis, Denis Zorin, Joan Bruna, Sanja Fidler, Or Litany
HIGHLIGHT: We present Neural Kernel Fields: a novel method for reconstructing implicit 3D shapes based on a learned kernel ridge regression.

1304, TITLE: A Deeper Dive Into What Deep Spatiotemporal Networks Encode: Quantifying Static vs. Dynamic Information

http://openaccess.thecvf.com//content/CVPR2022/html/Kowal_A_Deeper_Dive_Into_What_Deep_Spatiotemporal_Networks_Encode
_Quantifying_CVPR_2022_paper.html
AUTHORS: Matthew Kowal, Mennatullah Siam, Md Amirul Islam, Neil D. B. Bruce, Richard P. Wildes, Konstantinos G.
Derpanis
HIGHLIGHT: For example, while it has been observed that action recognition algorithms are heavily influenced by visual
appearance in single static frames, there is no quantitative methodology for evaluating such static bias in the latent representation
compared to bias toward dynamic information (e.g., motion). We tackle this challenge by proposing a novel approach for quantifying
the static and dynamic biases of any spatiotemporal model.

1305, TITLE: Detector-Free Weakly Supervised Group Activity Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Kim_Detector-
Free_Weakly_Supervised_Group_Activity_Recognition_CVPR_2022_paper.html
AUTHORS: Dongkeun Kim, Jinsung Lee, Minsu Cho, Suha Kwak
HIGHLIGHT: Existing models for this task are often impractical in that they demand ground-truth bounding box labels of
actors even in testing or rely on off-the-shelf object detectors. Motivated by this, we propose a novel model for group activity
recognition that depends neither on bounding box labels nor on object detector.

1306, TITLE: NFormer: Robust Person Re-Identification With Neighbor Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_NFormer_Robust_Person_Re-
Identification_With_Neighbor_Transformer_CVPR_2022_paper.html
AUTHORS: Haochen Wang, Jiayi Shen, Yongtuo Liu, Yan Gao, Efstratios Gavves
HIGHLIGHT: However, due to the high intra-identity variations, ignoring such interactions typically leads to outlier features.
To tackle this issue, we propose a Neighbor Transformer Network, or NFormer, which explicitly models interactions across all input
images, thus suppressing outlier features and leading to more robust representations overall.

1307, TITLE: Joint Forecasting of Panoptic Segmentations With Difference Attention
http://openaccess.thecvf.com//content/CVPR2022/html/Graber_Joint_Forecasting_of_Panoptic_Segmentations_With_Difference_Atte
ntion_CVPR_2022_paper.html
AUTHORS: Colin Graber, Cyril Jazra, Wenjie Luo, Liangyan Gui, Alexander G. Schwing
HIGHLIGHT: To address both issues, we study a new panoptic segmentation forecasting model that jointly forecasts all object
instances in a scene using a transformer model based on 'difference attention.'

1308, TITLE: HairCLIP: Design Your Hair by Text and Reference Image
http://openaccess.thecvf.com//content/CVPR2022/html/Wei_HairCLIP_Design_Your_Hair_by_Text_and_Reference_Image_CVPR_
2022_paper.html
AUTHORS: Tianyi Wei, Dongdong Chen, Wenbo Zhou, Jing Liao, Zhentao Tan, Lu Yuan, Weiming Zhang, Nenghai Yu
HIGHLIGHT: For this purpose, we encode the image and text conditions in a shared embedding space and propose a unified
hair editing framework by leveraging the powerful image text representation capability of the Contrastive Language-Image Pre-
Training (CLIP) model.

1309, TITLE: Imposing Consistency for Optical Flow Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Jeong_Imposing_Consistency_for_Optical_Flow_Estimation_CVPR_2022_pa
per.html
AUTHORS: Jisoo Jeong, Jamie Menjay Lin, Fatih Porikli, Nojun Kwak
HIGHLIGHT: This paper introduces novel and effective consistency strategies for optical flow estimation, a problem where
labels from real-world data are very challenging to derive.

1310, TITLE: Style Transformer for Image Inversion and Editing
http://openaccess.thecvf.com//content/CVPR2022/html/Hu_Style_Transformer_for_Image_Inversion_and_Editing_CVPR_2022_pape
r.html
AUTHORS: Xueqi Hu, Qiusheng Huang, Zhengyi Shi, Siyuan Li, Changxin Gao, Li Sun, Qingli Li
HIGHLIGHT: This paper presents a transformer-based image inversion and editing model for pretrained StyleGAN which is
not only with less distortions, but also of high quality and flexibility for editing.

1311, TITLE: OakInk: A Large-Scale Knowledge Repository for Understanding Hand-Object Interaction
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_OakInk_A_Large-
Scale_Knowledge_Repository_for_Understanding_Hand-Object_Interaction_CVPR_2022_paper.html
AUTHORS: Lixin Yang, Kailin Li, Xinyu Zhan, Fei Wu, Anran Xu, Liu Liu, Cewu Lu
HIGHLIGHT: In this work, we propose a multi-modal and rich-annotated knowledge repository, OakInk, for visual and
cognitive understanding of hand-object interactions.

1312, TITLE: Pyramid Adversarial Training Improves ViT Performance
http://openaccess.thecvf.com//content/CVPR2022/html/Herrmann_Pyramid_Adversarial_Training_Improves_ViT_Performance_CVP
R_2022_paper.html
AUTHORS: Charles Herrmann, Kyle Sargent, Lu Jiang, Ramin Zabih, Huiwen Chang, Ce Liu, Dilip Krishnan, Deqing Sun
HIGHLIGHT: In this work, we present pyramid adversarial training (PyramidAT), a simple and effective technique to improve
ViT's overall performance.

1313, TITLE: Bridging Global Context Interactions for High-Fidelity Image Completion

http://openaccess.thecvf.com//content/CVPR2022/html/Zheng_Bridging_Global_Context_Interactions_for_High-Fidelity_Image_Completion_CVPR_2022_paper.html
AUTHORS: Chuanxia Zheng, Tat-Jen Cham, Jianfei Cai, Dinh Phung
HIGHLIGHT: In this paper, we propose to treat image completion as a directionless sequence-to-sequence prediction task, and deploy a transformer to directly capture long-range dependence.

1314, TITLE: SwinBERT: End-to-End Transformers With Sparse Attention for Video Captioning
http://openaccess.thecvf.com//content/CVPR2022/html/Lin_SwinBERT_End-to-End_Transformers_With_Sparse_Attention_for_Video_Captioning_CVPR_2022_paper.html
AUTHORS: Kevin Lin, Linjie Li, Chung-Ching Lin, Faisal Ahmed, Zhe Gan, Zicheng Liu, Yumao Lu, Lijuan Wang
HIGHLIGHT: In this work, we present SwinBERT, an end-to-end transformer-based model for video captioning, which takes video frame patches directly as inputs, and outputs a natural language description.

1315, TITLE: Maximum Spatial Perturbation Consistency for Unpaired Image-to-Image Translation
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_Maximum_Spatial_Perturbation_Consistency_for_Unpaired_Image-to-Image_Translation_CVPR_2022_paper.html
AUTHORS: Yanwu Xu, Shaoan Xie, Wenhao Wu, Kun Zhang, Mingming Gong, Kayhan Batmanghelich
HIGHLIGHT: This paper proposes a universal regularization technique called maximum spatial perturbation consistency (MSPC), which enforces a spatial perturbation function (T) and the translation operator (G) to be commutative (i.e., $T \circ G = G \circ T$).

1316, TITLE: Unseen Classes at a Later Time? No Problem
http://openaccess.thecvf.com//content/CVPR2022/html/Kuchibhotla_Unseen_Classes_at_a_Later_Time_No_Problem_CVPR_2022_paper.html
AUTHORS: Hari Chandana Kuchibhotla, Sumitra S Malagi, Shivam Chandhok, Vineeth N Balasubramanian
HIGHLIGHT: A few recent efforts towards tackling CGZSL have been limited by difference in settings, practicality, data splits and protocols followed - inhibiting fair comparison and a clear direction forward. Motivated from these observations, in this work, we firstly consolidate the different CGZSL setting variants and propose a new Online CGZSL setting which is more practical and flexible.

1317, TITLE: InfoNeRF: Ray Entropy Minimization for Few-Shot Neural Volume Rendering
http://openaccess.thecvf.com//content/CVPR2022/html/Kim_InfoNeRF_Ray_Entropy_Minimization_for_Few-Shot_Neural_Volume_Rendering_CVPR_2022_paper.html
AUTHORS: Mijeong Kim, Seonguk Seo, Bohyung Han
HIGHLIGHT: We present an information-theoretic regularization technique for few-shot novel view synthesis based on neural implicit representation.

1318, TITLE: Learning the Degradation Distribution for Blind Image Super-Resolution
http://openaccess.thecvf.com//content/CVPR2022/html/Luo_Learning_the_Degradation_Distribution_for_Blind_Image_Super-Resolution_CVPR_2022_paper.html
AUTHORS: Zhengxiong Luo, Yan Huang, Shang Li, Liang Wang, Tieniu Tan
HIGHLIGHT: In this paper, we propose a probabilistic degradation model (PDM), which studies the degradation D as a random variable, and learns its distribution by modeling the mapping from a priori random variable z to D.

1319, TITLE: Dist-PU: Positive-Unlabeled Learning From a Label Distribution Perspective
http://openaccess.thecvf.com//content/CVPR2022/html/Zhao_Dist-PU_Positive-Unlabeled_Learning_From_a_Label_Distribution_Perspective_CVPR_2022_paper.html
AUTHORS: Yunrui Zhao, Qianqian Xu, Yangbangyan Jiang, Peisong Wen, Qingming Huang
HIGHLIGHT: While existing cost-sensitive-based methods have achieved state-of-the-art performances, they explicitly minimize the risk of classifying unlabeled data as negative samples, which might result in a negative-prediction preference of the classifier. To alleviate this issue, we resort to a label distribution perspective for PU learning in this paper.

1320, TITLE: SC2-PCR: A Second Order Spatial Compatibility for Efficient and Robust Point Cloud Registration
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_SC2-PCR_A_Second_Order_Spatial_Compatibility_for_Efficient_and_Robust_CVPR_2022_paper.html
AUTHORS: Zhi Chen, Kun Sun, Fan Yang, Wenbing Tao
HIGHLIGHT: In this paper, we present a second order spatial compatibility (SC^2) measure based method for efficient and robust point cloud registration (PCR), called SC^2-PCR.

1321, TITLE: Relative Pose From a Calibrated and an Uncalibrated Smartphone Image
http://openaccess.thecvf.com//content/CVPR2022/html/Ding_Relative_Pose_From_a_Calibrated_and_an_Uncalibrated_Smartphone_Image_CVPR_2022_paper.html
AUTHORS: Yaqing Ding, Daniel Barath, Jian Yang, Zuzana Kukelova
HIGHLIGHT: In this paper, we propose a new minimal and a non-minimal solver for estimating the relative camera pose together with the unknown focal length of the second camera.

1322, TITLE: Towards Robust and Reproducible Active Learning Using Neural Networks
http://openaccess.thecvf.com//content/CVPR2022/html/Munjal_Towards_Robust_and_Reproducible_Active_Learning_Using_Neural_Networks_CVPR_2022_paper.html
AUTHORS: Prateek Munjal, Nasir Hayat, Munawar Hayat, Jamshid Sourati, Shadab Khan

HIGHLIGHT: In this study, we demonstrate that under identical experimental conditions, different types of AL algorithms (uncertainty based, diversity based, and committee based) produce an inconsistent gain over random sampling baseline.

1323, TITLE: Retrieval Augmented Classification for Long-Tail Visual Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Long_Retrieval_Augmented_Classification_for_Long-Tail_Visual_Recognition_CVPR_2022_paper.html
AUTHORS: Alexander Long, Wei Yin, Thalaiyasingam Ajanthan, Vu Nguyen, Pulak Purkait, Ravi Garg, Alan Blair, Chunhua Shen, Anton van den Hengel
HIGHLIGHT: We introduce Retrieval Augmented Classification (RAC), a generic approach to augmenting standard image classification pipelines with an explicit retrieval module.

1324, TITLE: Not All Tokens Are Equal: Human-Centric Visual Analysis via Token Clustering Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Zeng_Not_All_Tokens_Are_Equal_Human-Centric_Visual_Analysis_via_Token_CVPR_2022_paper.html
AUTHORS: Wang Zeng, Sheng Jin, Wentao Liu, Chen Qian, Ping Luo, Wanli Ouyang, Xiaogang Wang
HIGHLIGHT: However, not all regions are equally important in human-centric vision tasks, e.g., the human body needs a fine representation with many tokens, while the image background can be modeled by a few tokens. To address this problem, we propose a novel Vision Transformer, called Token Clustering Transformer (TCFormer), which merges tokens by progressive clustering, where the tokens can be merged from different locations with flexible shapes and sizes.

1325, TITLE: Temporally Efficient Vision Transformer for Video Instance Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Temporally_Efficient_Vision_Transformer_for_Video_Instance_Segmentation_CVPR_2022_paper.html
AUTHORS: Shusheng Yang, Xinggang Wang, Yu Li, Yuxin Fang, Jiemin Fang, Wenyu Liu, Xun Zhao, Ying Shan
HIGHLIGHT: To effectively and efficiently model the crucial temporal information within a video clip, we propose a Temporally Efficient Vision Transformer (TeViT) for video instance segmentation (VIS).

1326, TITLE: The Devil Is in the Margin: Margin-Based Label Smoothing for Network Calibration
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_The_Devil_Is_in_the_Margin_Margin-Based_Label_Smoothing_for_CVPR_2022_paper.html
AUTHORS: Bingyuan Liu, Ismail Ben Ayed, Adrian Galdran, Jose Dolz
HIGHLIGHT: Following our observations, we propose a simple and flexible generalization based on inequality constraints, which imposes a controllable margin on logit distances.

1327, TITLE: NLX-GPT: A Model for Natural Language Explanations in Vision and Vision-Language Tasks
http://openaccess.thecvf.com//content/CVPR2022/html/Sammani_NLX-GPT_A_Model_for_Natural_Language_Explanations_in_Vision_and_CVPR_2022_paper.html
AUTHORS: Fawaz Sammani, Tanmoy Mukherjee, Nikos Deligiannis
HIGHLIGHT: We introduce NLX-GPT, a general, compact and faithful language model that can simultaneously predict an answer and explain it.

1328, TITLE: Bringing Old Films Back to Life
http://openaccess.thecvf.com//content/CVPR2022/html/Wan_Bringing_Old_Films_Back_to_Life_CVPR_2022_paper.html
AUTHORS: Ziyu Wan, Bo Zhang, Dongdong Chen, Jing Liao
HIGHLIGHT: We present a learning-based framework, recurrent transformer network (RTN), to restore heavily degraded old films.

1329, TITLE: Sound and Visual Representation Learning With Multiple Pretraining Tasks
http://openaccess.thecvf.com//content/CVPR2022/html/Vasudevan_Sound_and_Visual_Representation_Learning_With_Multiple_Pretraining_Tasks_CVPR_2022_paper.html
AUTHORS: Arun Balajee Vasudevan, Dengxin Dai, Luc Van Gool
HIGHLIGHT: The learned feature representations can exhibit different performance for each downstream task. In this light, this work aims to combine Multiple SSL tasks (Multi-SSL) that generalizes well for all downstream tasks.

1330, TITLE: WarpingGAN: Warping Multiple Uniform Priors for Adversarial 3D Point Cloud Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Tang_WarpingGAN_Warping_Multiple_Uniform_Priors_for_Adversarial_3D_Point_Cloud_CVPR_2022_paper.html
AUTHORS: Yingzhi Tang, Yue Qian, Qijian Zhang, Yiming Zeng, Junhui Hou, Xuefei Zhe
HIGHLIGHT: We propose WarpingGAN, an effective and efficient 3D point cloud generation network.

1331, TITLE: RePaint: Inpainting Using Denoising Diffusion Probabilistic Models
http://openaccess.thecvf.com//content/CVPR2022/html/Lugmayr_RePaint_Inpainting_Using_Denoising_Diffusion_Probabilistic_Models_CVPR_2022_paper.html
AUTHORS: Andreas Lugmayr, Martin Danelljan, Andres Romero, Fisher Yu, Radu Timofte, Luc Van Gool
HIGHLIGHT: In this work, we propose RePaint: A Denoising Diffusion Probabilistic Model (DDPM) based inpainting approach that is applicable to even extreme masks.

1332, TITLE: Revealing Occlusions With 4D Neural Fields
http://openaccess.thecvf.com//content/CVPR2022/html/Van_Hoorick_Revealing_Occlusions_With_4D_Neural_Fields_CVPR_2022_paper.html

AUTHORS: Basile Van Hoorick, Purva Tendulkar, Dídac Surís, Dennis Park, Simon Stent, Carl Vondrick
HIGHLIGHT: We introduce a framework for learning to estimate 4D visual representations from monocular RGB-D video, which is able to persist objects, even once they become obstructed by occlusions.

1333, TITLE: Meta Agent Teaming Active Learning for Pose Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Gong_Meta_Agent_Teaming_Active_Learning_for_Pose_Estimation_CVPR_2022_paper.html
AUTHORS: Jia Gong, Zhipeng Fan, Qiuhong Ke, Hossein Rahmani, Jun Liu
HIGHLIGHT: To reduce the human efforts on pose annotations, we propose a novel Meta Agent Teaming Active Learning (MATAL) framework to actively select and label informative images for effective learning.

1334, TITLE: Forward Propagation, Backward Regression, and Pose Association for Hand Tracking in the Wild
http://openaccess.thecvf.com//content/CVPR2022/html/Huang_Forward_Propagation_Backward_Regression_and_Pose_Association_for_Hand_Tracking_CVPR_2022_paper.html
AUTHORS: Mingzhen Huang, Supreeth Narasimhaswamy, Saif Vazir, Haibin Ling, Minh Hoai
HIGHLIGHT: We propose HandLer, a novel convolutional architecture that can jointly detect and track hands online in unconstrained videos.

1335, TITLE: Pseudo-Q: Generating Pseudo Language Queries for Visual Grounding
http://openaccess.thecvf.com//content/CVPR2022/html/Jiang_Pseudo-Q_Generating_Pseudo_Language_Queries_for_Visual_Grounding_CVPR_2022_paper.html
AUTHORS: Haojun Jiang, Yuanze Lin, Dongchen Han, Shiji Song, Gao Huang
HIGHLIGHT: To eliminate the heavy dependence on human annotations, we present a novel method, named Pseudo-Q, to automatically generate pseudo language queries for supervised training.

1336, TITLE: E2(GO)MOTION: Motion Augmented Event Stream for Egocentric Action Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Plizzari_E2GOMOTION_Motion_Augmented_Event_Stream_for_Egocentric_Action_Recognition_CVPR_2022_paper.html
AUTHORS: Chiara Plizzari, Mirco Planamente, Gabriele Goletto, Marco Cannici, Emanuele Gusso, Matteo Matteucci, Barbara Caputo
HIGHLIGHT: In this paper, we show that event data is a very valuable modality for egocentric action recognition.

1337, TITLE: ES6D: A Computation Efficient and Symmetry-Aware 6D Pose Regression Framework
http://openaccess.thecvf.com//content/CVPR2022/html/Mo_ES6D_A_Computation_Efficient_and_Symmetry-Aware_6D_Pose_Regression_Framework_CVPR_2022_paper.html
AUTHORS: Ningkai Mo, Wanshui Gan, Naoto Yokoya, Shifeng Chen
HIGHLIGHT: In this paper, a computation efficient regression framework is presented for estimating the 6D pose of rigid objects from a single RGB-D image, which is applicable to handling symmetric objects.

1338, TITLE: Self-Supervised Deep Image Restoration via Adaptive Stochastic Gradient Langevin Dynamics
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Self-Supervised_Deep_Image_Restoration_via_Adaptive_Stochastic_Gradient_Langevin_Dynamics_CVPR_2022_paper.html
AUTHORS: Weixi Wang, Ji Li, Hui Ji
HIGHLIGHT: Based on the neuralization of a Bayesian estimator of the problem, this paper presents a self-supervised deep learning approach to general image restoration problems.

1339, TITLE: Towards Discovering the Effectiveness of Moderately Confident Samples for Semi-Supervised Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Tang_Towards_Discovering_the_Effectiveness_of_Moderately_Confident_Samples_for_Semi-Supervised_CVPR_2022_paper.html
AUTHORS: Hui Tang, Kui Jia
HIGHLIGHT: In this work, we study whether the moderately confident samples are useless and how to select the useful ones to improve model optimization.

1340, TITLE: OoD-Bench: Quantifying and Understanding Two Dimensions of Out-of-Distribution Generalization
http://openaccess.thecvf.com//content/CVPR2022/html/Ye_OoD-Bench_Quantifying_and_Understanding_Two_Dimensions_of_Out-of-Distribution_Generalization_CVPR_2022_paper.html
AUTHORS: Nanyang Ye, Kaican Li, Haoyue Bai, Runpeng Yu, Lanqing Hong, Fengwei Zhou, Zhenguo Li, Jun Zhu
HIGHLIGHT: In this work, we first identify and measure two distinct kinds of distribution shifts that are ubiquitous in various datasets. Next, through extensive experiments, we compare OoD generalization algorithms across two groups of benchmarks, each dominated by one of the distribution shifts, revealing their strengths on one shift as well as limitations on the other shift.

1341, TITLE: An Empirical Study of Training End-to-End Vision-and-Language Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Dou_An_Empirical_Study_of_Training_End-to-End_Vision-and-Language_Transformers_CVPR_2022_paper.html
AUTHORS: Zi-Yi Dou, Yichong Xu, Zhe Gan, Jianfeng Wang, Shuohang Wang, Lijuan Wang, Chenguang Zhu, Pengchuan Zhang, Lu Yuan, Nanyun Peng, Zicheng Liu, Michael Zeng
HIGHLIGHT: In this paper, we present METER, a Multimodal End-to-end TransformER framework, through which we investigate how to design and pre-train a fully transformer-based VL model in an end-to-end manner.

1342, TITLE: Multimodal Dynamics: Dynamical Fusion for Trustworthy Multimodal Classification

142

http://openaccess.thecvf.com//content/CVPR2022/html/Han_Multimodal_Dynamics_Dynamical_Fusion_for_Trustworthy_Multimodal_Classification_CVPR_2022_paper.html
AUTHORS: Zongbo Han, Fan Yang, Junzhou Huang, Changqing Zhang, Jianhua Yao
HIGHLIGHT: However, conventional approaches are basically weak in providing trustworthy multimodal fusion, especially for safety-critical applications (e.g., medical diagnosis). For this issue, we propose a novel trustworthy multimodal classification algorithm termed Multimodal Dynamics, which dynamically evaluates both the feature-level and modality-level informativeness for different samples and thus trustworthily integrates multiple modalities.

1343, TITLE: The Neurally-Guided Shape Parser: Grammar-Based Labeling of 3D Shape Regions With Approximate Inference
http://openaccess.thecvf.com//content/CVPR2022/html/Jones_The_Neurally-Guided_Shape_Parser_Grammar-Based_Labeling_of_3D_Shape_Regions_CVPR_2022_paper.html
AUTHORS: R. Kenny Jones, Aalia Habib, Rana Hanocka, Daniel Ritchie
HIGHLIGHT: We propose the Neurally-Guided Shape Parser (NGSP), a method that learns how to assign fine-grained semantic labels to regions of a 3D shape.

1344, TITLE: Unsupervised Homography Estimation With Coplanarity-Aware GAN
http://openaccess.thecvf.com//content/CVPR2022/html/Hong_Unsupervised_Homography_Estimation_With_Coplanarity-Aware_GAN_CVPR_2022_paper.html
AUTHORS: Mingbo Hong, Yuhang Lu, Nianjin Ye, Chunyu Lin, Qijun Zhao, Shuaicheng Liu
HIGHLIGHT: In this work, we propose a novel method HomoGAN to guide unsupervised homography estimation to focus on the dominant plane.

1345, TITLE: LIFT: Learning 4D LiDAR Image Fusion Transformer for 3D Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Zeng_LIFT_Learning_4D_LiDAR_Image_Fusion_Transformer_for_3D_Object_CVPR_2022_paper.html
AUTHORS: Yihan Zeng, Da Zhang, Chunwei Wang, Zhenwei Miao, Ting Liu, Xin Zhan, Dayang Hao, Chao Ma
HIGHLIGHT: In this paper, we propose a novel LiDAR Image Fusion Transformer (LIFT) to model the mutual interaction relationship of cross-sensor data over time.

1346, TITLE: AutoLoss-Zero: Searching Loss Functions From Scratch for Generic Tasks
http://openaccess.thecvf.com//content/CVPR2022/html/Li_AutoLoss-Zero_Searching_Loss_Functions_From_Scratch_for_Generic_Tasks_CVPR_2022_paper.html
AUTHORS: Hao Li, Tianwen Fu, Jifeng Dai, Hongsheng Li, Gao Huang, Xizhou Zhu
HIGHLIGHT: In this paper, we propose AutoLoss-Zero, which is a general framework for searching loss functions from scratch for generic tasks.

1347, TITLE: PatchNet: A Simple Face Anti-Spoofing Framework via Fine-Grained Patch Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_PatchNet_A_Simple_Face_Anti-Spoofing_Framework_via_Fine-Grained_Patch_Recognition_CVPR_2022_paper.html
AUTHORS: Chien-Yi Wang, Yu-Ding Lu, Shang-Ta Yang, Shang-Hong Lai
HIGHLIGHT: In this work, we propose PatchNet which reformulates face anti-spoofing as a fine-grained patch-type recognition problem.

1348, TITLE: OnePose: One-Shot Object Pose Estimation Without CAD Models
http://openaccess.thecvf.com//content/CVPR2022/html/Sun_OnePose_One-Shot_Object_Pose_Estimation_Without_CAD_Models_CVPR_2022_paper.html
AUTHORS: Jiaming Sun, Zihao Wang, Siyu Zhang, Xingyi He, Hongcheng Zhao, Guofeng Zhang, Xiaowei Zhou
HIGHLIGHT: We propose a new method named OnePose for object pose estimation.

1349, TITLE: Weakly-Supervised Online Action Segmentation in Multi-View Instructional Videos
http://openaccess.thecvf.com//content/CVPR2022/html/Ghoddoosian_Weakly-Supervised_Online_Action_Segmentation_in_Multi-View_Instructional_Videos_CVPR_2022_paper.html
AUTHORS: Reza Ghoddoosian, Isht Dwivedi, Nakul Agarwal, Chiho Choi, Behzad Dariush
HIGHLIGHT: This paper addresses a new problem of weakly-supervised online action segmentation in instructional videos. We present a framework to segment streaming videos online at test time using Dynamic Programming and show its advantages over greedy sliding window approach.

1350, TITLE: Rethinking Minimal Sufficient Representation in Contrastive Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Rethinking_Minimal_Sufficient_Representation_in_Contrastive_Learning_CVPR_2022_paper.html
AUTHORS: Haoqing Wang, Xun Guo, Zhi-Hong Deng, Yan Lu
HIGHLIGHT: This reveals a new problem that the contrastive learning models have the risk of over-fitting to the shared information between views. To alleviate this problem, we propose to increase the mutual information between the representation and input as regularization to approximately introduce more task-relevant information, since we cannot utilize any downstream task information during training.

1351, TITLE: Disentangling Visual Embeddings for Attributes and Objects
http://openaccess.thecvf.com//content/CVPR2022/html/Saini_Disentangling_Visual_Embeddings_for_Attributes_and_Objects_CVPR_2022_paper.html

AUTHORS: Nirat Saini, Khoi Pham, Abhinav Shrivastava
HIGHLIGHT: We study the problem of compositional zero-shot learning for object-attribute recognition.

1352, TITLE: Scalable Penalized Regression for Noise Detection in Learning With Noisy Labels
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Scalable_Penalized_Regression_for_Noise_Detection_in_Learning_With_Noisy_CVPR_2022_paper.html
AUTHORS: Yikai Wang, Xinwei Sun, Yanwei Fu
HIGHLIGHT: In this paper, we propose using a theoretically guaranteed noisy label detection framework to detect and remove noisy data for Learning with Noisy Labels (LNL).

1353, TITLE: Effective Conditioned and Composed Image Retrieval Combining CLIP-Based Features
http://openaccess.thecvf.com//content/CVPR2022/html/Baldrati_Effective_Conditioned_and_Composed_Image_Retrieval_Combining_CLIP-Based_Features_CVPR_2022_paper.html
AUTHORS: Alberto Baldrati, Marco Bertini, Tiberio Uricchio, Alberto Del Bimbo
HIGHLIGHT: In this demo, we present an interactive system based on a combiner network, trained using contrastive learning, that combines visual and textual features obtained from the OpenAI CLIP network to address conditioned CBIR.

1354, TITLE: Registering Explicit to Implicit: Towards High-Fidelity Garment Mesh Reconstruction From Single Images
http://openaccess.thecvf.com//content/CVPR2022/html/Zhu_Registering_Explicit_to_Implicit_Towards_High-Fidelity_Garment_Mesh_Reconstruction_From_CVPR_2022_paper.html
AUTHORS: Heming Zhu, Lingteng Qiu, Yuda Qiu, Xiaoguang Han
HIGHLIGHT: However, a common problem for the implicit-based methods is that they cannot produce separated and topology-consistent mesh for each garment piece, which is crucial for the current 3D content creation pipeline. To address this issue, we proposed a novel geometry inference framework ReEF that reconstructs topology- consistent layered garment mesh by registering the explicit garment template to the whole-body implicit fields predicted from single images.

1355, TITLE: Federated Class-Incremental Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Dong_Federated_Class-Incremental_Learning_CVPR_2022_paper.html
AUTHORS: Jiahua Dong, Lixu Wang, Zhen Fang, Gan Sun, Shichao Xu, Xiao Wang, Qi Zhu
HIGHLIGHT: To tackle the global forgetting brought by the non-i.i.d class imbalance across clients, we propose a proxy server that selects the best old global model to assist the local relation distillation.

1356, TITLE: MiniViT: Compressing Vision Transformers With Weight Multiplexing
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_MiniViT_Compressing_Vision_Transformers_With_Weight_Multiplexing_CVPR_2022_paper.html
AUTHORS: Jinnian Zhang, Houwen Peng, Kan Wu, Mengchen Liu, Bin Xiao, Jianlong Fu, Lu Yuan
HIGHLIGHT: However, ViT models suffer from huge number of parameters, restricting their applicability on devices with limited computation. To alleviate this problem, we propose MiniViT, a new compression framework, which achieves parameter reduction in vision transformers while retaining the same performance.

1357, TITLE: Practical Stereo Matching via Cascaded Recurrent Network With Adaptive Correlation
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Practical_Stereo_Matching_via_Cascaded_Recurrent_Network_With_Adaptive_Correlation_CVPR_2022_paper.html
AUTHORS: Jiankun Li, Peisen Wang, Pengfei Xiong, Tao Cai, Ziwei Yan, Lei Yang, Jiangyu Liu, Haoqiang Fan, Shuaicheng Liu
HIGHLIGHT: In this paper, we propose a set of innovative designs to tackle the problem of practical stereo matching: 1) to better recover fine depth details, we design a hierarchical network with recurrent refinement to update disparities in a coarse-to-fine manner, as well as a stacked cascaded architecture for inference; 2) we propose an adaptive group correlation layer to mitigate the impact of erroneous rectification; 3) we introduce a new synthetic dataset with special attention to difficult cases for better generalizing to real-world scenes.

1358, TITLE: D-Grasp: Physically Plausible Dynamic Grasp Synthesis for Hand-Object Interactions
http://openaccess.thecvf.com//content/CVPR2022/html/Christen_D-Grasp_Physically_Plausible_Dynamic_Grasp_Synthesis_for_Hand-Object_Interactions_CVPR_2022_paper.html
AUTHORS: Sammy Christen, Muhammed Kocabas, Emre Aksan, Jemin Hwangbo, Jie Song, Otmar Hilliges
HIGHLIGHT: We introduce the dynamic grasp synthesis task: given an object with a known 6D pose and a grasp reference, our goal is to generate motions that move the object to a target 6D pose.

1359, TITLE: Show, Deconfound and Tell: Image Captioning With Causal Inference
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Show_Deconfound_and_Tell_Image_Captioning_With_Causal_Inference_CVPR_2022_paper.html
AUTHORS: Bing Liu, Dong Wang, Xu Yang, Yong Zhou, Rui Yao, Zhiwen Shao, Jiaqi Zhao
HIGHLIGHT: In this paper, we first use Structural Causal Models (SCMs) to show how two confounders damage the image captioning. Then we apply the backdoor adjustment to propose a novel causal inference based image captioning (CIIC) framework, which consists of an interventional object detector (IOD) and an interventional transformer decoder (ITD) to jointly confront both confounders.

1360, TITLE: Extracting Triangular 3D Models, Materials, and Lighting From Images
http://openaccess.thecvf.com//content/CVPR2022/html/Munkberg_Extracting_Triangular_3D_Models_Materials_and_Lighting_From_Images_CVPR_2022_paper.html

AUTHORS: Jacob Munkberg, Jon Hasselgren, Tianchang Shen, Jun Gao, Wenzheng Chen, Alex Evans, Thomas Müller, Sanja Fidler

HIGHLIGHT: We present an efficient method for joint optimization of topology, materials and lighting from multi-view image observations.

1361, TITLE: Weakly Supervised Segmentation on Outdoor 4D Point Clouds With Temporal Matching and Spatial Graph Propagation

http://openaccess.thecvf.com//content/CVPR2022/html/Shi_Weakly_Supervised_Segmentation_on_Outdoor_4D_Point_Clouds_With _Temporal_CVPR_2022_paper.html

AUTHORS: Hanyu Shi, Jiacheng Wei, Ruibo Li, Fayao Liu, Guosheng Lin

HIGHLIGHT: In this paper, we study how to achieve scene understanding with limited annotated data.

1362, TITLE: ImFace: A Nonlinear 3D Morphable Face Model With Implicit Neural Representations

http://openaccess.thecvf.com//content/CVPR2022/html/Zheng_ImFace_A_Nonlinear_3D_Morphable_Face_Model_With_Implicit_N eural_CVPR_2022_paper.html

AUTHORS: Mingwu Zheng, Hongyu Yang, Di Huang, Liming Chen

HIGHLIGHT: This paper presents a novel 3D morphable face model, namely ImFace, to learn a nonlinear and continuous space with implicit neural representations.

1363, TITLE: MobRecon: Mobile-Friendly Hand Mesh Reconstruction From Monocular Image

http://openaccess.thecvf.com//content/CVPR2022/html/Chen_MobRecon_Mobile-Friendly_Hand_Mesh_Reconstruction_From_Monocular_Image_CVPR_2022_paper.html

AUTHORS: Xingyu Chen, Yufeng Liu, Yajiao Dong, Xiong Zhang, Chongyang Ma, Yanmin Xiong, Yuan Zhang, Xiaoyan Guo

HIGHLIGHT: In this work, we propose a framework for single-view hand mesh reconstruction, which can simultaneously achieve high reconstruction accuracy, fast inference speed, and temporal coherence.

1364, TITLE: Layered Depth Refinement With Mask Guidance

http://openaccess.thecvf.com//content/CVPR2022/html/Kim_Layered_Depth_Refinement_With_Mask_Guidance_CVPR_2022_paper .html

AUTHORS: Soo Ye Kim, Jianming Zhang, Simon Niklaus, Yifei Fan, Simon Chen, Zhe Lin, Munchurl Kim

HIGHLIGHT: Hence, in this paper, we formulate a novel problem of mask-guided depth refinement that utilizes a generic mask to refine the depth prediction of SIDE models.

1365, TITLE: Parameter-Free Online Test-Time Adaptation

http://openaccess.thecvf.com//content/CVPR2022/html/Boudiaf_Parameter-Free_Online_Test-Time_Adaptation_CVPR_2022_paper.html

AUTHORS: Malik Boudiaf, Romain Mueller, Ismail Ben Ayed, Luca Bertinetto

HIGHLIGHT: Motivated by the inherent uncertainty around the conditions that will ultimately be encountered at test time, we propose a particularly "conservative" approach, which addresses the problem with a Laplacian Adjusted Maximum-likelihood Estimation (LAME) objective.

1366, TITLE: SIGMA: Semantic-Complete Graph Matching for Domain Adaptive Object Detection

http://openaccess.thecvf.com//content/CVPR2022/html/Li_SIGMA_Semantic-Complete_Graph_Matching_for_Domain_Adaptive_Object_Detection_CVPR_2022_paper.html

AUTHORS: Wuyang Li, Xinyu Liu, Yixuan Yuan

HIGHLIGHT: Though great success, they ignore the significant within-class variance and the domain-mismatched semantics within the training batch, leading to a sub-optimal adaptation. To overcome these challenges, we propose a novel SemantIc-complete Graph MAtching (SIGMA) framework for DAOD, which completes mismatched semantics and reformulates the adaptation with graph matching.

1367, TITLE: Global Convergence of MAML and Theory-Inspired Neural Architecture Search for Few-Shot Learning

http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Global_Convergence_of_MAML_and_Theory-Inspired_Neural_Architecture_Search_for_CVPR_2022_paper.html

AUTHORS: Haoxiang Wang, Yite Wang, Ruoyu Sun, Bo Li

HIGHLIGHT: In this paper, we first prove that MAML with over-parameterized DNNs is guaranteed to converge to global optima at a linear rate. Our convergence analysis indicates that MAML with over-parameterized DNNs is equivalent to kernel regression with a novel class of kernels, which we name as Meta Neural Tangent Kernels (MetaNTK). Then, we propose MetaNTK-NAS, a new training-free neural architecture search (NAS) method for few-shot learning that uses MetaNTK to rank and select architectures.

1368, TITLE: LAKe-Net: Topology-Aware Point Cloud Completion by Localizing Aligned Keypoints

http://openaccess.thecvf.com//content/CVPR2022/html/Tang_LAKe-Net_Topology-Aware_Point_Cloud_Completion_by_Localizing_Aligned_Keypoints_CVPR_2022_paper.html

AUTHORS: Junshu Tang, Zhijun Gong, Ran Yi, Yuan Xie, Lizhuang Ma

HIGHLIGHT: To better tackle the missing topology part, we propose LAKe-Net, a novel topology-aware point cloud completion model by localizing aligned keypoints, with a novel Keypoints-Skeleton-Shape prediction manner.

1369, TITLE: Scribble-Supervised LiDAR Semantic Segmentation

http://openaccess.thecvf.com//content/CVPR2022/html/Unal_Scribble-Supervised_LiDAR_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Ozan Unal, Dengxin Dai, Luc Van Gool
HIGHLIGHT: In this paper, we propose using scribbles to annotate LiDAR point clouds and release ScribbleKITTI, the first scribble-annotated dataset for LiDAR semantic segmentation.

1370, TITLE: AlignMixup: Improving Representations by Interpolating Aligned Features
http://openaccess.thecvf.com//content/CVPR2022/html/Venkataramanan_AlignMixup_Improving_Representations_by_Interpolating_Aligned_Features_CVPR_2022_paper.html
AUTHORS: Shashanka Venkataramanan, Ewa Kijak, Laurent Amsaleg, Yannis Avrithis
HIGHLIGHT: In this work, we revisit mixup from the deformation perspective and introduce AlignMixup, where we geometrically align two images in the feature space.

1371, TITLE: No Pain, Big Gain: Classify Dynamic Point Cloud Sequences With Static Models by Fitting Feature-Level Space-Time Surfaces
http://openaccess.thecvf.com//content/CVPR2022/html/Zhong_No_Pain_Big_Gain_Classify_Dynamic_Point_Cloud_Sequences_With_CVPR_2022_paper.html
AUTHORS: Jia-Xing Zhong, Kaichen Zhou, Qingyong Hu, Bing Wang, Niki Trigoni, Andrew Markham
HIGHLIGHT: To capture 3D motions without explicitly tracking correspondences, we propose a kinematics-inspired neural network (Kinet) by generalizing the kinematic concept of ST-surfaces to the feature space.

1372, TITLE: HiVT: Hierarchical Vector Transformer for Multi-Agent Motion Prediction
http://openaccess.thecvf.com//content/CVPR2022/html/Zhou_HiVT_Hierarchical_Vector_Transformer_for_Multi-Agent_Motion_Prediction_CVPR_2022_paper.html
AUTHORS: Zikang Zhou, Luyao Ye, Jianping Wang, Kui Wu, Kejie Lu
HIGHLIGHT: However, existing methods neglect the symmetries of the problem and suffer from the expensive computational cost, facing the challenge of making real-time multi-agent motion prediction without sacrificing the prediction performance. To tackle this challenge, we propose Hierarchical Vector Transformer (HiVT) for fast and accurate multi-agent motion prediction.

1373, TITLE: HerosNet: Hyperspectral Explicable Reconstruction and Optimal Sampling Deep Network for Snapshot Compressive Imaging
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_HerosNet_Hyperspectral_Explicable_Reconstruction_and_Optimal_Sampling_Deep_Network_for_CVPR_2022_paper.html
AUTHORS: Xuanyu Zhang, Yongbing Zhang, Ruiqin Xiong, Qilin Sun, Jian Zhang
HIGHLIGHT: In this paper, we propose a novel Hyperspectral Explicable Reconstruction and Optimal Sampling deep Network for SCI, dubbed HerosNet, which includes several phases under the ISTA-unfolding framework.

1374, TITLE: Vision Transformer Slimming: Multi-Dimension Searching in Continuous Optimization Space
http://openaccess.thecvf.com//content/CVPR2022/html/Chavan_Vision_Transformer_Slimming_Multi-Dimension_Searching_in_Continuous_Optimization_Space_CVPR_2022_paper.html
AUTHORS: Arnav Chavan, Zhiqiang Shen, Zhuang Liu, Zechun Liu, Kwang-Ting Cheng, Eric P. Xing
HIGHLIGHT: This paper explores the feasibility of finding an optimal sub-model from a vision transformer and introduces a pure vision transformer slimming (ViT-Slim) framework.

1375, TITLE: Brain-Inspired Multilayer Perceptron With Spiking Neurons
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Brain-Inspired_Multilayer_Perceptron_With_Spiking_Neurons_CVPR_2022_paper.html
AUTHORS: Wenshuo Li, Hanting Chen, Jianyuan Guo, Ziyang Zhang, Yunhe Wang
HIGHLIGHT: In this paper, we incorporate the machanism of LIF neurons into the MLP models, to achieve better accuracy without extra FLOPs.

1376, TITLE: Learning To Estimate Robust 3D Human Mesh From In-the-Wild Crowded Scenes
http://openaccess.thecvf.com//content/CVPR2022/html/Choi_Learning_To_Estimate_Robust_3D_Human_Mesh_From_In-the-Wild_Crowded_CVPR_2022_paper.html
AUTHORS: Hongsuk Choi, Gyeongsik Moon, JoonKyu Park, Kyoung Mu Lee
HIGHLIGHT: We consider the problem of recovering a single person's 3D human mesh from in-the-wild crowded scenes.

1377, TITLE: ObjectFormer for Image Manipulation Detection and Localization
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_ObjectFormer_for_Image_Manipulation_Detection_and_Localization_CVPR_2022_paper.html
AUTHORS: Junke Wang, Zuxuan Wu, Jingjing Chen, Xintong Han, Abhinav Shrivastava, Ser-Nam Lim, Yu-Gang Jiang
HIGHLIGHT: In this paper, we propose ObjectFormer to detect and localize image manipulations.

1378, TITLE: Detecting Deepfakes With Self-Blended Images
http://openaccess.thecvf.com//content/CVPR2022/html/Shiohara_Detecting_Deepfakes_With_Self-Blended_Images_CVPR_2022_paper.html
AUTHORS: Kaede Shiohara, Toshihiko Yamasaki
HIGHLIGHT: In this paper, we present novel synthetic training data called self-blended images (SBIs) to detect deepfakes.

1379, TITLE: Correlation-Aware Deep Tracking

http://openaccess.thecvf.com//content/CVPR2022/html/Xie_Correlation-Aware_Deep_Tracking_CVPR_2022_paper.html
AUTHORS: Fei Xie, Chunyu Wang, Guangting Wang, Yue Cao, Wankou Yang, Wenjun Zeng
HIGHLIGHT: While most methods focus on designing robust matching operations, we propose a novel target-dependent
feature network inspired by the self-/cross-attention scheme.

1380, TITLE: Learnable Irrelevant Modality Dropout for Multimodal Action Recognition on Modality-Specific Annotated
Videos
http://openaccess.thecvf.com//content/CVPR2022/html/Alfasly_Learnable_Irrelevant_Modality_Dropout_for_Multimodal_Action_Re
cognition_on_Modality-Specific_CVPR_2022_paper.html
AUTHORS: Saghir Alfasly, Jian Lu, Chen Xu, Yuru Zou
HIGHLIGHT: However, effectively leveraging the audio modality in vision-specific annotated videos for action recognition is
of particular challenge. To tackle this challenge, we propose a novel audio-visual framework that effectively leverages the audio
modality in any solely vision-specific annotated dataset.

1381, TITLE: NeurMiPs: Neural Mixture of Planar Experts for View Synthesis
http://openaccess.thecvf.com//content/CVPR2022/html/Lin_NeurMiPs_Neural_Mixture_of_Planar_Experts_for_View_Synthesis_CV
PR_2022_paper.html
AUTHORS: Zhi-Hao Lin, Wei-Chiu Ma, Hao-Yu Hsu, Yu-Chiang Frank Wang, Shenlong Wang
HIGHLIGHT: We present Neural Mixtures of Planar Experts (NeurMiPs), a novel planar-based scene representation for
modeling geometry and appearance.

1382, TITLE: Implicit Sample Extension for Unsupervised Person Re-Identification
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Implicit_Sample_Extension_for_Unsupervised_Person_Re-
Identification_CVPR_2022_paper.html
AUTHORS: Xinyu Zhang, Dongdong Li, Zhigang Wang, Jian Wang, Errui Ding, Javen Qinfeng Shi, Zhaoxiang Zhang,
Jingdong Wang
HIGHLIGHT: Due to the limited samples in each identity, we suppose there may lack some underlying information to well
reveal the accurate clusters. To discover these information, we propose an Implicit Sample Extension (ISE) method to generate what
we call support samples around the cluster boundaries.

1383, TITLE: Energy-Based Latent Aligner for Incremental Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Joseph_Energy-
Based_Latent_Aligner_for_Incremental_Learning_CVPR_2022_paper.html
AUTHORS: K J Joseph, Salman Khan, Fahad Shahbaz Khan, Rao Muhammad Anwer, Vineeth N Balasubramanian
HIGHLIGHT: In this work, we propose ELI: Energy-based Latent Aligner for Incremental Learning, which first learns an
energy manifold for the latent representations such that previous task latents will have low energy and the current task latents have
high energy values.

1384, TITLE: Towards Semi-Supervised Deep Facial Expression Recognition With an Adaptive Confidence Margin
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Towards_Semi-
Supervised_Deep_Facial_Expression_Recognition_With_an_Adaptive_Confidence_CVPR_2022_paper.html
AUTHORS: Hangyu Li, Nannan Wang, Xi Yang, Xiaoyu Wang, Xinbo Gao
HIGHLIGHT: In this paper, we learn an Adaptive Confidence Margin (Ada-CM) to fully leverage all unlabeled data for semi-
supervised deep facial expression recognition.

1385, TITLE: GanOrCon: Are Generative Models Useful for Few-Shot Segmentation?
http://openaccess.thecvf.com//content/CVPR2022/html/Saha_GanOrCon_Are_Generative_Models_Useful_for_Few-
Shot_Segmentation_CVPR_2022_paper.html
AUTHORS: Oindrila Saha, Zezhou Cheng, Subhransu Maji
HIGHLIGHT: But how do these improvements stack-up against recent advances in self-supervised learning? Motivated by this
we present an alternative approach based on contrastive learning and compare their performance on standard few-shot part
segmentation benchmarks.

1386, TITLE: Bi-Level Doubly Variational Learning for Energy-Based Latent Variable Models
http://openaccess.thecvf.com//content/CVPR2022/html/Kan_Bi-Level_Doubly_Variational_Learning_for_Energy-
Based_Latent_Variable_Models_CVPR_2022_paper.html
AUTHORS: Ge Kan, Jinhu Lü, Tian Wang, Baochang Zhang, Aichun Zhu, Lei Huang, Guodong Guo, Hichem
Snoussi
HIGHLIGHT: In this paper, we propose Bi-level doubly variational learning (BiDVL), which is based on a new bi-level
optimization framework and two tractable variational distributions to facilitate learning EBLVMs.

1387, TITLE: SplitNets: Designing Neural Architectures for Efficient Distributed Computing on Head-Mounted Systems
http://openaccess.thecvf.com//content/CVPR2022/html/Dong_SplitNets_Designing_Neural_Architectures_for_Efficient_Distributed_
Computing_on_Head-Mounted_CVPR_2022_paper.html
AUTHORS: Xin Dong, Barbara De Salvo, Meng Li, Chiao Liu, Zhongnan Qu, H.T. Kung, Ziyun Li
HIGHLIGHT: To achieve an optimal balance among computation, communication, and performance, a split-aware neural
architecture search framework, SplitNets, is introduced to conduct model designing, splitting, and communication reduction
simultaneously.

1388, TITLE: Masked-Attention Mask Transformer for Universal Image Segmentation

http://openaccess.thecvf.com//content/CVPR2022/html/Cheng_Masked-Attention_Mask_Transformer_for_Universal_Image_Segmentation_CVPR_2022_paper.html
AUTHORS: Bowen Cheng, Ishan Misra, Alexander G. Schwing, Alexander Kirillov, Rohit Girdhar
HIGHLIGHT: We present Masked-attention Mask Transformer (Mask2Former), a new architecture capable of addressing any image segmentation task (panoptic, instance or semantic).

1389, TITLE: Reading To Listen at the Cocktail Party: Multi-Modal Speech Separation
http://openaccess.thecvf.com//content/CVPR2022/html/Rahimi_Reading_To_Listen_at_the_Cocktail_Party_Multi-Modal_Speech_Separation_CVPR_2022_paper.html
AUTHORS: Akam Rahimi, Triantafyllos Afouras, Andrew Zisserman
HIGHLIGHT: The goal of this paper is speech separation and enhancement in multi-speaker and noisy environments using a combination of different modalities.

1390, TITLE: AxIoU: An Axiomatically Justified Measure for Video Moment Retrieval
http://openaccess.thecvf.com//content/CVPR2022/html/Togashi_AxIoU_An_Axiomatically_Justified_Measure_for_Video_Moment_Retrieval_CVPR_2022_paper.html
AUTHORS: Riku Togashi, Mayu Otani, Yuta Nakashima, Esa Rahtu, Janne Heikkilä, Tetsuya Sakai
HIGHLIGHT: First, it is rank-insensitive: It ignores the rank positions of successfully localised moments in the top-K ranked list by treating the list as a set. Second, it binarises the Intersection over Union (IoU) of each retrieved video moment using the threshold \theta and thereby ignoring fine-grained localisation quality of ranked moments. We propose an alternative measure for evaluating VMR, called Average Max IoU (AxIoU), which is free from the above two problems.

1391, TITLE: NOC-REK: Novel Object Captioning With Retrieved Vocabulary From External Knowledge
http://openaccess.thecvf.com//content/CVPR2022/html/Vo_NOC-REK_Novel_Object_Captioning_With_Retrieved_Vocabulary_From_External_Knowledge_CVPR_2022_paper.html
AUTHORS: Duc Minh Vo, Hong Chen, Akihiro Sugimoto, Hideki Nakayama
HIGHLIGHT: We propose an end-to-end Novel Object Captioning with Retrieved vocabulary from External Knowledge method (NOC-REK), which simultaneously learns vocabulary retrieval and caption generation, successfully describing novel objects outside of the training dataset.

1392, TITLE: Boosting Robustness of Image Matting With Context Assembling and Strong Data Augmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Dai_Boosting_Robustness_of_Image_Matting_With_Context_Assembling_and_Strong_CVPR_2022_paper.html
AUTHORS: Yutong Dai, Brian Price, He Zhang, Chunhua Shen
HIGHLIGHT: Although some works propose to either refine the trimaps or adapt the algorithms to real-world images via extra data augmentation, none of them has taken both into consideration, not to mention the significant performance deterioration on benchmarks while using those data augmentation. To fill this gap, we propose an image matting method which achieves higher robustness (RMat) via multilevel context assembling and strong data augmentation targeting matting.

1393, TITLE: Group R-CNN for Weakly Semi-Supervised Object Detection With Points
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Group_R-CNN_for_Weakly_Semi-Supervised_Object_Detection_With_Points_CVPR_2022_paper.html
AUTHORS: Shilong Zhang, Zhuoran Yu, Liyang Liu, Xinjiang Wang, Aojun Zhou, Kai Chen
HIGHLIGHT: Based on the classic R-CNN architecture, we propose an effective point-to-box regressor: Group R-CNN.

1394, TITLE: Weakly-Supervised Action Transition Learning for Stochastic Human Motion Prediction
http://openaccess.thecvf.com//content/CVPR2022/html/Mao_Weakly-Supervised_Action_Transition_Learning_for_Stochastic_Human_Motion_Prediction_CVPR_2022_paper.html
AUTHORS: Wei Mao, Miaomiao Liu, Mathieu Salzmann
HIGHLIGHT: We introduce the task of action-driven stochastic human motion prediction, which aims to predict multiple plausible future motions given a sequence of action labels and a short motion history.

1395, TITLE: Speech Driven Tongue Animation
http://openaccess.thecvf.com//content/CVPR2022/html/Medina_Speech_Driven_Tongue_Animation_CVPR_2022_paper.html
AUTHORS: Salvador Medina, Denis Tome, Carsten Stoll, Mark Tiede, Kevin Munhall, Alexander G. Hauptmann, Iain Matthews
HIGHLIGHT: In this work, we introduce a large-scale speech and mocap dataset that focuses on capturing tongue, jaw, and lip motion.

1396, TITLE: Hybrid Relation Guided Set Matching for Few-Shot Action Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Hybrid_Relation_Guided_Set_Matching_for_Few-Shot_Action_Recognition_CVPR_2022_paper.html
AUTHORS: Xiang Wang, Shiwei Zhang, Zhiwu Qing, Mingqian Tang, Zhengrong Zuo, Changxin Gao, Rong Jin, Nong Sang
HIGHLIGHT: To overcome the two limitations, we propose a novel Hybrid Relation guided Set Matching (HyRSM) approach that incorporates two key components: hybrid relation module and set matching metric.

1397, TITLE: Self-Supervised Spatial Reasoning on Multi-View Line Drawings
http://openaccess.thecvf.com//content/CVPR2022/html/Xiang_Self-Supervised_Spatial_Reasoning_on_Multi-View_Line_Drawings_CVPR_2022_paper.html

AUTHORS: Siyuan Xiang, Anbang Yang, Yanfei Xue, Yaoqing Yang, Chen Feng
HIGHLIGHT: Based on the fact that self-supervised learning is helpful when a large number of data are available, we propose two self-supervised learning approaches to improve the baseline performance for view consistency reasoning and camera pose reasoning tasks on the SPARE3D dataset.

1398, TITLE: Language-Bridged Spatial-Temporal Interaction for Referring Video Object Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Ding_Language-Bridged_Spatial-Temporal_Interaction_for_Referring_Video_Object_Segmentation_CVPR_2022_paper.html
AUTHORS: Zihan Ding, Tianrui Hui, Junshi Huang, Xiaoming Wei, Jizhong Han, Si Liu
HIGHLIGHT: However, these methods suffer from spatial misalignment or false distractors due to delayed and implicit spatial-temporal interaction occurring in the decoding phase. To tackle these limitations, we propose a Language-Bridged Duplex Transfer (LBDT) module which utilizes language as an intermediary bridge to accomplish explicit and adaptive spatial-temporal interaction earlier in the encoding phase.

1399, TITLE: Cross-Patch Dense Contrastive Learning for Semi-Supervised Segmentation of Cellular Nuclei in Histopathologic Images
http://openaccess.thecvf.com//content/CVPR2022/html/Wu_Cross-Patch_Dense_Contrastive_Learning_for_Semi-Supervised_Segmentation_of_Cellular_Nuclei_CVPR_2022_paper.html
AUTHORS: Huisi Wu, Zhaoze Wang, Youyi Song, Lin Yang, Jing Qin
HIGHLIGHT: We study the semi-supervised learning problem, using a few labeled data and a large amount of unlabeled data to train the network, by developing a cross-patch dense contrastive learning framework, to segment cellular nuclei in histopathologic images.

1400, TITLE: Frame-Wise Action Representations for Long Videos via Sequence Contrastive Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Frame-Wise_Action_Representations_for_Long_Videos_via_Sequence_Contrastive_Learning_CVPR_2022_paper.html
AUTHORS: Minghao Chen, Fangyun Wei, Chong Li, Deng Cai
HIGHLIGHT: In this paper, we introduce a novel contrastive action representation learning (CARL) framework to learn frame-wise action representations, especially for long videos, in a self-supervised manner.

1401, TITLE: Coarse-To-Fine Deep Video Coding With Hyperprior-Guided Mode Prediction
http://openaccess.thecvf.com//content/CVPR2022/html/Hu_Coarse-To-Fine_Deep_Video_Coding_With_Hyperprior-Guided_Mode_Prediction_CVPR_2022_paper.html
AUTHORS: Zhihao Hu, Guo Lu, Jinyang Guo, Shan Liu, Wei Jiang, Dong Xu
HIGHLIGHT: Specifically, using hyperprior information as the input, we propose two mode prediction networks to respectively predict the optimal block resolutions for better motion coding and decide whether to skip residual information from each block for better residual coding without introducing additional bit cost while bringing negligible extra computation cost.

1402, TITLE: Generalized Binary Search Network for Highly-Efficient Multi-View Stereo
http://openaccess.thecvf.com//content/CVPR2022/html/Mi_Generalized_Binary_Search_Network_for_Highly-Efficient_Multi-View_Stereo_CVPR_2022_paper.html
AUTHORS: Zhenxing Mi, Chang Di, Dan Xu
HIGHLIGHT: In this work, we propose a novel method for highly efficient MVS that remarkably decreases the memory footprint, meanwhile clearly advancing state-of-the-art depth prediction performance.

1403, TITLE: SHIFT: A Synthetic Driving Dataset for Continuous Multi-Task Domain Adaptation
http://openaccess.thecvf.com//content/CVPR2022/html/Sun_SHIFT_A_Synthetic_Driving_Dataset_for_Continuous_Multi-Task_Domain_Adaptation_CVPR_2022_paper.html
AUTHORS: Tao Sun, Mattia Segu, Janis Postels, Yuxuan Wang, Luc Van Gool, Bernt Schiele, Federico Tombari, Fisher Yu
HIGHLIGHT: In this paper, we introduce the largest synthetic dataset for autonomous driving, SHIFT.

1404, TITLE: Adaptive Hierarchical Representation Learning for Long-Tailed Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Adaptive_Hierarchical_Representation_Learning_for_Long-Tailed_Object_Detection_CVPR_2022_paper.html
AUTHORS: Banghuai Li
HIGHLIGHT: In this paper, we propose a novel method, named Adaptive Hierarchical Representation Learning (AHRL), from a metric learning perspective to address long-tailed object detection.

1405, TITLE: FlexIT: Towards Flexible Semantic Image Translation
http://openaccess.thecvf.com//content/CVPR2022/html/Couairon_FlexIT_Towards_Flexible_Semantic_Image_Translation_CVPR_2022_paper.html
AUTHORS: Guillaume Couairon, Asya Grechka, Jakob Verbeek, Holger Schwenk, Matthieu Cord
HIGHLIGHT: We propose FlexIT, a novel method which can take any input image and a user-defined text instruction for editing.

1406, TITLE: Face2Exp: Combating Data Biases for Facial Expression Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Zeng_Face2Exp_Combating_Data_Biases_for_Facial_Expression_Recognition_CVPR_2022_paper.html
AUTHORS: Dan Zeng, Zhiyuan Lin, Xiao Yan, Yuting Liu, Fei Wang, Bo Tang

149

HIGHLIGHT: To combat the mismatch, we propose the Meta-Face2Exp framework, which consists of a base network and an adaptation network.

1407, TITLE: SAR-Net: Shape Alignment and Recovery Network for Category-Level 6D Object Pose and Size Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Lin_SAR-Net_Shape_Alignment_and_Recovery_Network_for_Category-Level_6D_Object_CVPR_2022_paper.html
AUTHORS: Haitao Lin, Zichang Liu, Chilam Cheang, Yanwei Fu, Guodong Guo, Xiangyang Xue
HIGHLIGHT: Given a single scene image, this paper proposes a method of Category-level 6D Object Pose and Size Estimation (COPSE) from the point cloud of the target object, without external real pose-annotated training data.

1408, TITLE: Whose Hands Are These? Hand Detection and Hand-Body Association in the Wild
http://openaccess.thecvf.com//content/CVPR2022/html/Narasimhaswamy_Whose_Hands_Are_These_Hand_Detection_and_Hand-Body_Association_in_CVPR_2022_paper.html
AUTHORS: Supreeth Narasimhaswamy, Thanh Nguyen, Mingzhen Huang, Minh Hoai
HIGHLIGHT: We propose a novel end-to-end trainable convolutional network that can jointly detect hands and the body location for the corresponding person.

1409, TITLE: Mega-NERF: Scalable Construction of Large-Scale NeRFs for Virtual Fly-Throughs
http://openaccess.thecvf.com//content/CVPR2022/html/Turki_Mega-NERF_Scalable_Construction_of_Large-Scale_NeRFs_for_Virtual_Fly-Throughs_CVPR_2022_paper.html
AUTHORS: Haithem Turki, Deva Ramanan, Mahadev Satyanarayanan
HIGHLIGHT: We introduce a simple geometric clustering algorithm for data parallelism that partitions training images (or rather pixels) into different NeRF submodules that can be trained in parallel.

1410, TITLE: PINA: Learning a Personalized Implicit Neural Avatar From a Single RGB-D Video Sequence
http://openaccess.thecvf.com//content/CVPR2022/html/Dong_PINA_Learning_a_Personalized_Implicit_Neural_Avatar_From_a_Single_CVPR_2022_paper.html
AUTHORS: Zijian Dong, Chen Guo, Jie Song, Xu Chen, Andreas Geiger, Otmar Hilliges
HIGHLIGHT: We present a novel method to learn Personalized Implicit Neural Avatars (PINA) from a short RGB-D sequence.

1411, TITLE: Forecasting From LiDAR via Future Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Peri_Forecasting_From_LiDAR_via_Future_Object_Detection_CVPR_2022_paper.html
AUTHORS: Neehar Peri, Jonathon Luiten, Mengtian Li, Aljoša Ošep, Laura Leal-Taixé, Deva Ramanan
HIGHLIGHT: In this paper, we propose an end-to-end approach for motion forecasting based on raw sensor measurement as opposed to ground truth tracks.

1412, TITLE: CRAFT: Cross-Attentional Flow Transformer for Robust Optical Flow
http://openaccess.thecvf.com//content/CVPR2022/html/Sui_CRAFT_Cross-Attentional_Flow_Transformer_for_Robust_Optical_Flow_CVPR_2022_paper.html
AUTHORS: Xiuchao Sui, Shaohua Li, Xue Geng, Yan Wu, Xinxing Xu, Yong Liu, Rick Goh, Hongyuan Zhu
HIGHLIGHT: On large displacements with motion blur, noisy correlations could cause severe errors in the estimated flow. To overcome this challenge, we propose a new architecture "CRoss-Attentional Flow Transformer" (CRAFT), aiming to revitalize the correlation volume computation.

1413, TITLE: Adversarial Eigen Attack on Black-Box Models
http://openaccess.thecvf.com//content/CVPR2022/html/Zhou_Adversarial_Eigen_Attack_on_Black-Box_Models_CVPR_2022_paper.html
AUTHORS: Linjun Zhou, Peng Cui, Xingxuan Zhang, Yinan Jiang, Shiqiang Yang
HIGHLIGHT: In this paper, we deal with a more practical setting where a pre-trained white-box model with network parameters is provided without extra training data.

1414, TITLE: Training Quantised Neural Networks With STE Variants: The Additive Noise Annealing Algorithm
http://openaccess.thecvf.com//content/CVPR2022/html/Spallanzani_Training_Quantised_Neural_Networks_With_STE_Variants_The_Additive_Noise_CVPR_2022_paper.html
AUTHORS: Matteo Spallanzani, Gian Paolo Leonardi, Luca Benini
HIGHLIGHT: In this paper, we analyse STE variants and study their impact on QNN training.

1415, TITLE: Split Hierarchical Variational Compression
http://openaccess.thecvf.com//content/CVPR2022/html/Ryder_Split_Hierarchical_Variational_Compression_CVPR_2022_paper.html
AUTHORS: Tom Ryder, Chen Zhang, Ning Kang, Shifeng Zhang
HIGHLIGHT: That is, not only do state-of-the-art methods, such as normalizing flows, often demonstrate out-performance, but the initial bits required in coding makes single and parallel image compression challenging. To remedy this, we introduce Split Hierarchical Variational Compression (SHVC).

1416, TITLE: Video Swin Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Video_Swin_Transformer_CVPR_2022_paper.html
AUTHORS: Ze Liu, Jia Ning, Yue Cao, Yixuan Wei, Zheng Zhang, Stephen Lin, Han Hu

HIGHLIGHT: In this paper, we instead advocate an inductive bias of locality in video Transformers, which leads to a better speed-accuracy trade-off compared to previous approaches which compute self-attention globally even with spatial-temporal factorization.

1417, TITLE: Privacy Preserving Partial Localization
http://openaccess.thecvf.com//content/CVPR2022/html/Geppert_Privacy_Preserving_Partial_Localization_CVPR_2022_paper.html
AUTHORS: Marcel Geppert, Viktor Larsson, Johannes L. Schönberger, Marc Pollefeys
HIGHLIGHT: However, in many scenarios, the pose itself might be sensitive information. We propose a principled approach overcoming these limitations, based on two observations.

1418, TITLE: Cross-Modal Background Suppression for Audio-Visual Event Localization
http://openaccess.thecvf.com//content/CVPR2022/html/Xia_Cross-Modal_Background_Suppression_for_Audio-Visual_Event_Localization_CVPR_2022_paper.html
AUTHORS: Yan Xia, Zhou Zhao
HIGHLIGHT: Hence this paper proposes a novel cross-modal background suppression network for AVE task, operating at the time- and event-level, aiming to improve localization performance through suppressing asynchronous audiovisual background frames from the examined events and reducing redundant noise.

1419, TITLE: Mutual Quantization for Cross-Modal Search With Noisy Labels
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Mutual_Quantization_for_Cross-Modal_Search_With_Noisy_Labels_CVPR_2022_paper.html
AUTHORS: Erkun Yang, Dongren Yao, Tongliang Liu, Cheng Deng
HIGHLIGHT: To address the challenge, we present a general robust cross-modal hashing framework to correlate distinct modalities and combat noisy labels simultaneously.

1420, TITLE: Lagrange Motion Analysis and View Embeddings for Improved Gait Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Chai_Lagrange_Motion_Analysis_and_View_Embeddings_for_Improved_Gait_Recognition_CVPR_2022_paper.html
AUTHORS: Tianrui Chai, Annan Li, Shaoxiong Zhang, Zilong Li, Yunhong Wang
HIGHLIGHT: In this paper, we analyzed human walking using the Lagrange's equation and come to the conclusion that second-order information in the temporal dimension is necessary for identification.

1421, TITLE: SphereSR: 360deg Image Super-Resolution With Arbitrary Projection via Continuous Spherical Image Representation
http://openaccess.thecvf.com//content/CVPR2022/html/Yoon_SphereSR_360deg_Image_Super-Resolution_With_Arbitrary_Projection_via_Continuous_Spherical_CVPR_2022_paper.html
AUTHORS: Youngho Yoon, Inchul Chung, Lin Wang, Kuk-Jin Yoon
HIGHLIGHT: In this paper, we propose SphereSR, a novel framework to generate a continuous spherical image representation from an LR 360deg image, with the goal of predicting the RGB values at given spherical coordinates for superresolution with an arbitrary 360deg image projection.

1422, TITLE: Neural Mesh Simplification
http://openaccess.thecvf.com//content/CVPR2022/html/Potamias_Neural_Mesh_Simplification_CVPR_2022_paper.html
AUTHORS: Rolandos Alexandros Potamias, Stylianos Ploumpis, Stefanos Zafeiriou
HIGHLIGHT: In this work we attempt to tackle the novel task of learnable and differentiable mesh simplification.

1423, TITLE: Cloth-Changing Person Re-Identification From a Single Image With Gait Prediction and Regularization
http://openaccess.thecvf.com//content/CVPR2022/html/Jin_Cloth-Changing_Person_Re-Identification_From_a_Single_Image_With_Gait_Prediction_CVPR_2022_paper.html
AUTHORS: Xin Jin, Tianyu He, Kecheng Zheng, Zhiheng Yin, Xu Shen, Zhen Huang, Ruoyu Feng, Jianqiang Huang, Zhibo Chen, Xian-Sheng Hua
HIGHLIGHT: In this paper, we focus on handling well the CC-ReID problem under a more challenging setting, i.e., just from a single image, which enables an efficient and latency-free person identity matching for surveillance.

1424, TITLE: BoxeR: Box-Attention for 2D and 3D Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Nguyen_BoxeR_Box-Attention_for_2D_and_3D_Transformers_CVPR_2022_paper.html
AUTHORS: Duy-Kien Nguyen, Jihong Ju, Olaf Booij, Martin R. Oswald, Cees G. M. Snoek
HIGHLIGHT: In this paper, we propose a simple attention mechanism, we call Box-Attention.

1425, TITLE: Neural Architecture Search With Representation Mutual Information
http://openaccess.thecvf.com//content/CVPR2022/html/Zheng_Neural_Architecture_Search_With_Representation_Mutual_Information_CVPR_2022_paper.html
AUTHORS: Xiawu Zheng, Xiang Fei, Lei Zhang, Chenglin Wu, Fei Chao, Jianzhuang Liu, Wei Zeng, Yonghong Tian, Rongrong Ji
HIGHLIGHT: Existing strategies, such as employing standard training or performance predictor, often suffer from high computational complexity and low generality. To address this issue, we propose to rank architectures by Representation Mutual Information (RMI).

1426, TITLE: Deep Hyperspectral-Depth Reconstruction Using Single Color-Dot Projection

http://openaccess.thecvf.com//content/CVPR2022/html/Li_Deep_Hyperspectral-Depth_Reconstruction_Using_Single_Color-Dot_Projection_CVPR_2022_paper.html
AUTHORS: Chunyu Li, Yusuke Monno, Masatoshi Okutomi
HIGHLIGHT: In this paper, we propose a novel single-shot hyperspectral-depth reconstruction method using an off-the-shelf RGB camera and projector.

1427, TITLE: M3T: Three-Dimensional Medical Image Classifier Using Multi-Plane and Multi-Slice Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Jang_M3T_Three-Dimensional_Medical_Image_Classifier_Using_Multi-Plane_and_Multi-Slice_Transformer_CVPR_2022_paper.html
AUTHORS: Jinseong Jang, Dosik Hwang
HIGHLIGHT: In this study, we propose a three-dimensional Medical image classifier using Multi-plane and Multi-slice Transformer (M3T) network to classify Alzheimer's disease (AD) in 3D MRI images.

1428, TITLE: 3MASSIV: Multilingual, Multimodal and Multi-Aspect Dataset of Social Media Short Videos
http://openaccess.thecvf.com//content/CVPR2022/html/Gupta_3MASSIV_Multilingual_Multimodal_and_Multi-Aspect_Dataset_of_Social_Media_Short_CVPR_2022_paper.html
AUTHORS: Vikram Gupta, Trisha Mittal, Puneet Mathur, Vaibhav Mishra, Mayank Maheshwari, Aniket Bera, Debdoot Mukherjee, Dinesh Manocha
HIGHLIGHT: We present 3MASSIV, a multilingual, multimodal and multi-aspect, expertly-annotated dataset of diverse short videos extracted from a social media platform.

1429, TITLE: Can Neural Nets Learn the Same Model Twice? Investigating Reproducibility and Double Descent From the Decision Boundary Perspective
http://openaccess.thecvf.com//content/CVPR2022/html/Somepalli_Can_Neural_Nets_Learn_the_Same_Model_Twice_Investigating_Reproducibility_CVPR_2022_paper.html
AUTHORS: Gowthami Somepalli, Liam Fowl, Arpit Bansal, Ping Yeh-Chiang, Yehuda Dar, Richard Baraniuk, Micah Goldblum, Tom Goldstein
HIGHLIGHT: We discuss methods for visualizing neural network decision boundaries and decision regions.

1430, TITLE: Cross Domain Object Detection by Target-Perceived Dual Branch Distillation
http://openaccess.thecvf.com//content/CVPR2022/html/He_Cross_Domain_Object_Detection_by_Target-Perceived_Dual_Branch_Distillation_CVPR_2022_paper.html
AUTHORS: Mengzhe He, Yali Wang, Jiaxi Wu, Yiru Wang, Hanqing Li, Bo Li, Weihao Gan, Wei Wu, Yu Qiao
HIGHLIGHT: Existing approaches mainly focus on either of these two difficulties, even though they are closely coupled in cross domain object detection. To solve this problem, we propose a novel Target-perceived Dual-branch Distillation (TDD) framework.

1431, TITLE: A Proposal-Based Paradigm for Self-Supervised Sound Source Localization in Videos
http://openaccess.thecvf.com//content/CVPR2022/html/Xuan_A_Proposal-Based_Paradigm_for_Self-Supervised_Sound_Source_Localization_in_Videos_CVPR_2022_paper.html
AUTHORS: Hanyu Xuan, Zhiliang Wu, Jian Yang, Yan Yan, Xavier Alameda-Pineda
HIGHLIGHT: In this paper, we advocate a novel proposal-based paradigm that can directly perform semantic object-level localization, without any manual annotations.

1432, TITLE: Overcoming Catastrophic Forgetting in Incremental Object Detection via Elastic Response Distillation
http://openaccess.thecvf.com//content/CVPR2022/html/Feng_Overcoming_Catastrophic_Forgetting_in_Incremental_Object_Detection_via_Elastic_Response_CVPR_2022_paper.html
AUTHORS: Tao Feng, Mang Wang, Hangjie Yuan
HIGHLIGHT: In this paper, we propose a response-based incremental distillation method, dubbed Elastic Response Distillation (ERD), which focuses on elastically learning responses from the classification head and the regression head.

1433, TITLE: GroupNet: Multiscale Hypergraph Neural Networks for Trajectory Prediction With Relational Reasoning
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_GroupNet_Multiscale_Hypergraph_Neural_Networks_for_Trajectory_Prediction_With_Relational_CVPR_2022_paper.html
AUTHORS: Chenxin Xu, Maosen Li, Zhenyang Ni, Ya Zhang, Siheng Chen
HIGHLIGHT: To promote more comprehensive interaction modeling for relational reasoning, we propose GroupNet, a multiscale hypergraph neural network, which is novel in terms of both interaction capturing and representation learning.

1434, TITLE: Unbiased Subclass Regularization for Semi-Supervised Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Guan_Unbiased_Subclass_Regularization_for_Semi-Supervised_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Dayan Guan, Jiaxing Huang, Aoran Xiao, Shijian Lu
HIGHLIGHT: This paper presents an unbiased subclass regularization network (USRN) that alleviates the class imbalance issue by learning class-unbiased segmentation from balanced subclass distributions.

1435, TITLE: P3IV: Probabilistic Procedure Planning From Instructional Videos With Weak Supervision
http://openaccess.thecvf.com//content/CVPR2022/html/Zhao_P3IV_Probabilistic_Procedure_Planning_From_Instructional_Videos_With_Weak_Supervision_CVPR_2022_paper.html
AUTHORS: He Zhao, Isma Hadji, Nikita Dvornik, Konstantinos G. Derpanis, Richard P. Wildes, Allan D. Jepson
HIGHLIGHT: In this paper, we study the problem of procedure planning in instructional videos.

1436, TITLE:	Hierarchical Nearest Neighbor Graph Embedding for Efficient Dimensionality Reduction
http://openaccess.thecvf.com//content/CVPR2022/html/Sarfraz_Hierarchical_Nearest_Neighbor_Graph_Embedding_for_Efficient_Dimensionality_Reduction_CVPR_2022_paper.html
AUTHORS:	Saquib Sarfraz, Marios Koulakis, Constantin Seibold, Rainer Stiefelhagen
HIGHLIGHT:	We introduce a novel method based on a hierarchy built on 1-nearest neighbor graphs in the original space which is used to preserve the grouping properties of the data distribution on multiple levels.

1437, TITLE:	Coupled Iterative Refinement for 6D Multi-Object Pose Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Lipson_Coupled_Iterative_Refinement_for_6D_Multi-Object_Pose_Estimation_CVPR_2022_paper.html
AUTHORS:	Lahav Lipson, Zachary Teed, Ankit Goyal, Jia Deng
HIGHLIGHT:	We propose a new approach to 6D object pose estimation which consists of an end-to-end differentiable architecture that makes use of geometric knowledge.

1438, TITLE:	Multi-View Transformer for 3D Visual Grounding
http://openaccess.thecvf.com//content/CVPR2022/html/Huang_Multi-View_Transformer_for_3D_Visual_Grounding_CVPR_2022_paper.html
AUTHORS:	Shijia Huang, Yilun Chen, Jiaya Jia, Liwei Wang
HIGHLIGHT:	In this paper, we propose a Multi-View Transformer (MVT) for 3D visual grounding.

1439, TITLE:	Structured Sparse R-CNN for Direct Scene Graph Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Teng_Structured_Sparse_R-CNN_for_Direct_Scene_Graph_Generation_CVPR_2022_paper.html
AUTHORS:	Yao Teng, Limin Wang
HIGHLIGHT:	Instead, from a perspective on SGG as a direct set prediction, this paper presents a simple, sparse, and unified framework, termed as Structured Sparse R-CNN.

1440, TITLE:	Multi-Grained Spatio-Temporal Features Perceived Network for Event-Based Lip-Reading
http://openaccess.thecvf.com//content/CVPR2022/html/Tan_Multi-Grained_Spatio-Temporal_Features_Perceived_Network_for_Event-Based_Lip-Reading_CVPR_2022_paper.html
AUTHORS:	Ganchao Tan, Yang Wang, Han Han, Yang Cao, Feng Wu, Zheng-Jun Zha
HIGHLIGHT:	In this work, we introduce a novel type of sensing device, event cameras, for the task of ALR.

1441, TITLE:	Semi-Supervised Video Paragraph Grounding With Contrastive Encoder
http://openaccess.thecvf.com//content/CVPR2022/html/Jiang_Semi-Supervised_Video_Paragraph_Grounding_With_Contrastive_Encoder_CVPR_2022_paper.html
AUTHORS:	Xun Jiang, Xing Xu, Jingran Zhang, Fumin Shen, Zuo Cao, Heng Tao Shen
HIGHLIGHT:	However, we find the existing VPG methods may not perform well on context modeling and highly rely on video-paragraph annotations. To tackle this problem, we propose a novel VPG method termed Semi-supervised Video-Paragraph TRansformer (SVPTR), which can more effectively exploit contextual information in paragraphs and significantly reduce the dependency on annotated data.

1442, TITLE:	Continual Predictive Learning From Videos
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Continual_Predictive_Learning_From_Videos_CVPR_2022_paper.html
AUTHORS:	Geng Chen, Wendong Zhang, Han Lu, Siyu Gao, Yunbo Wang, Mingsheng Long, Xiaokang Yang
HIGHLIGHT:	Can we develop predictive learning algorithms that can deal with more realistic, non-stationary physical environments? In this paper, we study a new continual learning problem in the context of video prediction, and observe that most existing methods suffer from severe catastrophic forgetting in this setup.

1443, TITLE:	Weakly Paired Associative Learning for Sound and Image Representations via Bimodal Associative Memory
http://openaccess.thecvf.com//content/CVPR2022/html/Lee_Weakly_Paired_Associative_Learning_for_Sound_and_Image_Representations_via_CVPR_2022_paper.html
AUTHORS:	Sangmin Lee, Hyung-Il Kim, Yong Man Ro
HIGHLIGHT:	Based on the observation, we propose a new problem to deal with the weakly paired condition: How to boost a certain modal representation even by using other unpaired modal data.

1444, TITLE:	BARC: Learning To Regress 3D Dog Shape From Images by Exploiting Breed Information
http://openaccess.thecvf.com//content/CVPR2022/html/Ruegg_BARC_Learning_To_Regress_3D_Dog_Shape_From_Images_by_CVPR_2022_paper.html
AUTHORS:	Nadine Rüegg, Silvia Zuffi, Konrad Schindler, Michael J. Black
HIGHLIGHT:	Our goal is to recover the 3D shape and pose of dogs from a single image.

1445, TITLE:	Knowledge Distillation: A Good Teacher Is Patient and Consistent
http://openaccess.thecvf.com//content/CVPR2022/html/Beyer_Knowledge_Distillation_A_Good_Teacher_Is_Patient_and_Consistent_CVPR_2022_paper.html
AUTHORS:	Lucas Beyer, Xiaohua Zhai, Amélie Royer, Larisa Markeeva, Rohan Anil, Alexander Kolesnikov

HIGHLIGHT: There is a growing discrepancy in computer vision between large-scale models that achieve state-of-the-art performance and models that are affordable in practical applications. In this paper we address this issue and significantly bridge the gap between these two types of models.

1446, TITLE: PCA-Based Knowledge Distillation Towards Lightweight and Content-Style Balanced Photorealistic Style Transfer Models
http://openaccess.thecvf.com//content/CVPR2022/html/Chiu_PCA-Based_Knowledge_Distillation_Towards_Lightweight_and_Content-Style_Balanced_Photorealistic_Style_CVPR_2022_paper.html
AUTHORS: Tai-Yin Chiu, Danna Gurari
HIGHLIGHT: We introduce PCA-based knowledge distillation to distill lightweight models and show it is motivated by theory.

1447, TITLE: Frame Averaging for Equivariant Shape Space Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Atzmon_Frame_Averaging_for_Equivariant_Shape_Space_Learning_CVPR_2022_paper.html
AUTHORS: Matan Atzmon, Koki Nagano, Sanja Fidler, Sameh Khamis, Yaron Lipman
HIGHLIGHT: In this paper, we present a framework for incorporating equivariance in encoders and decoders by introducing two contributions: (i) adapting the recent Frame Averaging (FA) framework for building generic, efficient, and maximally expressive Equivariant autoencoders; and (ii) constructing autoencoders equivariant to piecewise Euclidean motions applied to different parts of the shape.

1448, TITLE: Transformer Tracking With Cyclic Shifting Window Attention
http://openaccess.thecvf.com//content/CVPR2022/html/Song_Transformer_Tracking_With_Cyclic_Shifting_Window_Attention_CVPR_2022_paper.html
AUTHORS: Zikai Song, Junqing Yu, Yi-Ping Phoebe Chen, Wei Yang
HIGHLIGHT: In this paper, we propose a new transformer architecture with multi-scale cyclic shifting window attention for visual object tracking, elevating the attention from pixel to window level.

1449, TITLE: ProposalCLIP: Unsupervised Open-Category Object Proposal Generation via Exploiting CLIP Cues
http://openaccess.thecvf.com//content/CVPR2022/html/Shi_ProposalCLIP_Unsupervised_Open-Category_Object_Proposal_Generation_via_Exploiting_CLIP_Cues_CVPR_2022_paper.html
AUTHORS: Hengcan Shi, Munawar Hayat, Yicheng Wu, Jianfei Cai
HIGHLIGHT: In this paper, we propose ProposalCLIP, a method towards unsupervised open-category object proposal generation.

1450, TITLE: Towards Understanding Adversarial Robustness of Optical Flow Networks
http://openaccess.thecvf.com//content/CVPR2022/html/Schrodi_Towards_Understanding_Adversarial_Robustness_of_Optical_Flow_Networks_CVPR_2022_paper.html
AUTHORS: Simon Schrodi, Tonmoy Saikia, Thomas Brox
HIGHLIGHT: In this paper, we analyze the cause of the problem and show that the lack of robustness is rooted in the classical aperture problem of optical flow estimation in combination with bad choices in the details of the network architecture.

1451, TITLE: Panoptic SegFormer: Delving Deeper Into Panoptic Segmentation With Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Panoptic_SegFormer_Delving_Deeper_Into_Panoptic_Segmentation_With_Transformers_CVPR_2022_paper.html
AUTHORS: Zhiqi Li, Wenhai Wang, Enze Xie, Zhiding Yu, Anima Anandkumar, Jose M. Alvarez, Ping Luo, Tong Lu
HIGHLIGHT: We present Panoptic SegFormer, a general framework for panoptic segmentation with transformers.

1452, TITLE: Training High-Performance Low-Latency Spiking Neural Networks by Differentiation on Spike Representation
http://openaccess.thecvf.com//content/CVPR2022/html/Meng_Training_High-Performance_Low-Latency_Spiking_Neural_Networks_by_Differentiation_on_Spike_CVPR_2022_paper.html
AUTHORS: Qingyan Meng, Mingqing Xiao, Shen Yan, Yisen Wang, Zhouchen Lin, Zhi-Quan Luo
HIGHLIGHT: In this paper, we propose the Differentiation on Spike Representation (DSR) method, which could achieve high performance that is competitive to ANNs yet with low latency.

1453, TITLE: AnyFace: Free-Style Text-To-Face Synthesis and Manipulation
http://openaccess.thecvf.com//content/CVPR2022/html/Sun_AnyFace_Free-Style_Text-To-Face_Synthesis_and_Manipulation_CVPR_2022_paper.html
AUTHORS: Jianxin Sun, Qiyao Deng, Qi Li, Muyi Sun, Min Ren, Zhenan Sun
HIGHLIGHT: So this paper proposes the first free-style text-to-face method namely AnyFace enabling much wider open world applications such as metaverse, social media, cosmetics, forensics, etc.

1454, TITLE: HL-Net: Heterophily Learning Network for Scene Graph Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Lin_HL-Net_Heterophily_Learning_Network_for_Scene_Graph_Generation_CVPR_2022_paper.html
AUTHORS: Xin Lin, Changxing Ding, Yibing Zhan, Zijian Li, Dacheng Tao
HIGHLIGHT: Accordingly, in this paper, we propose a novel Heterophily Learning Network (HL-Net) to comprehensively explore the homophily and heterophily between objects/relationships in scene graphs.

1455, TITLE: Lifelong Graph Learning

http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Lifelong_Graph_Learning_CVPR_2022_paper.html
AUTHORS: Chen Wang, Yuheng Qiu, Dasong Gao, Sebastian Scherer
HIGHLIGHT: In this paper, we bridge GNN and lifelong learning by converting a continual graph learning problem to a regular graph learning problem so GNN can inherit the lifelong learning techniques developed for convolutional neural networks (CNN).

1456, TITLE: Hypergraph-Induced Semantic Tuplet Loss for Deep Metric Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Lim_Hypergraph-Induced_Semantic_Tuplet_Loss_for_Deep_Metric_Learning_CVPR_2022_paper.html
AUTHORS: Jongin Lim, Sangdoo Yun, Seulki Park, Jin Young Choi
HIGHLIGHT: In this paper, we propose Hypergraph-Induced Semantic Tuplet (HIST) loss for deep metric learning that leverages the multilateral semantic relations of multiple samples to multiple classes via hypergraph modeling.

1457, TITLE: Computing Wasserstein-p Distance Between Images With Linear Cost
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Computing_Wasserstein-p_Distance_Between_Images_With_Linear_Cost_CVPR_2022_paper.html
AUTHORS: Yidong Chen, Chen Li, Zhonghua Lu
HIGHLIGHT: In this paper, we propose a novel algorithm to compute the Wasserstein-p distance between discrete measures by restricting the optimal transport (OT) problem on a subset.

1458, TITLE: DLFormer: Discrete Latent Transformer for Video Inpainting
http://openaccess.thecvf.com//content/CVPR2022/html/Ren_DLFormer_Discrete_Latent_Transformer_for_Video_Inpainting_CVPR_2022_paper.html
AUTHORS: Jingjing Ren, Qingqing Zheng, Yuanyuan Zhao, Xuemiao Xu, Chen Li
HIGHLIGHT: While noticing the capability of discrete representation for complex reasoning and predictive learning, we propose a novel Discrete Latent Transformer (DLFormer) to reformulate video inpainting tasks into the discrete latent space rather the previous continuous feature space.

1459, TITLE: Unsupervised Representation Learning for Binary Networks by Joint Classifier Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Kim_Unsupervised_Representation_Learning_for_Binary_Networks_by_Joint_Classifier_Learning_CVPR_2022_paper.html
AUTHORS: Dahyun Kim, Jonghyun Choi
HIGHLIGHT: But such networks are not readily deployable to edge devices. To accelerate deployment of models with the benefit of unsupervised representation learning to such resource limited devices for various downstream tasks, we propose a self-supervised learning method for binary networks that uses a moving target network.

1460, TITLE: High Quality Segmentation for Ultra High-Resolution Images
http://openaccess.thecvf.com//content/CVPR2022/html/Shen_High_Quality_Segmentation_for_Ultra_High-Resolution_Images_CVPR_2022_paper.html
AUTHORS: Tiancheng Shen, Yuechen Zhang, Lu Qi, Jason Kuen, Xingyu Xie, Jianlong Wu, Zhe Lin, Jiaya Jia
HIGHLIGHT: Motivated by the fact that humans distinguish among objects continuously from coarse to precise levels, we propose the Continuous Refinement Model(CRM) for the ultra high-resolution segmentation refinement task.

1461, TITLE: Investigating Tradeoffs in Real-World Video Super-Resolution
http://openaccess.thecvf.com//content/CVPR2022/html/Chan_Investigating_Tradeoffs_in_Real-World_Video_Super-Resolution_CVPR_2022_paper.html
AUTHORS: Kelvin C.K. Chan, Shangchen Zhou, Xiangyu Xu, Chen Change Loy
HIGHLIGHT: To facilitate fair comparisons, we propose the new VideoLQ dataset, which contains a large variety of real-world low-quality video sequences containing rich textures and patterns.

1462, TITLE: MERLOT Reserve: Neural Script Knowledge Through Vision and Language and Sound
http://openaccess.thecvf.com//content/CVPR2022/html/Zellers_MERLOT_Reserve_Neural_Script_Knowledge_Through_Vision_and_Language_and_CVPR_2022_paper.html
AUTHORS: Rowan Zellers, Jiasen Lu, Ximing Lu, Youngjae Yu, Yanpeng Zhao, Mohammadreza Salehi, Aditya Kusupati, Jack Hessel, Ali Farhadi, Yejin Choi
HIGHLIGHT: We introduce MERLOT Reserve, a model that represents videos jointly over time -- through a new training objective that learns from audio, subtitles, and video frames.

1463, TITLE: Differentiable Stereopsis: Meshes From Multiple Views Using Differentiable Rendering
http://openaccess.thecvf.com//content/CVPR2022/html/Goel_Differentiable_Stereopsis_Meshes_From_Multiple_Views_Using_Differentiable_Rendering_CVPR_2022_paper.html
AUTHORS: Shubham Goel, Georgia Gkioxari, Jitendra Malik
HIGHLIGHT: We propose Differentiable Stereopsis, a multi-view stereo approach that reconstructs shape and texture from few input views and noisy cameras.

1464, TITLE: Towards Practical Certifiable Patch Defense With Vision Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Towards_Practical_Certifiable_Patch_Defense_With_Vision_Transformer_CVPR_2022_paper.html
AUTHORS: Zhaoyu Chen, Bo Li, Jianghe Xu, Shuang Wu, Shouhong Ding, Wenqiang Zhang

HIGHLIGHT: To move towards a practical certifiable patch defense, we introduce Vision Transformer (ViT) into the framework of Derandomized Smoothing (DS).

1465, TITLE: A Conservative Approach for Unbiased Learning on Unknown Biases
http://openaccess.thecvf.com//content/CVPR2022/html/Jeon_A_Conservative_Approach_for_Unbiased_Learning_on_Unknown_Biases_CVPR_2022_paper.html
AUTHORS: Myeongho Jeon, Daekyung Kim, Woochul Lee, Myungjoo Kang, Joonseok Lee
HIGHLIGHT: To mitigate this issue, we present a new scenario that does not necessitate a predefined bias. Under the observation that CNNs do have multi-variant and unbiased representations in the model, we propose a conservative framework that employs this internal information for unbiased learning.

1466, TITLE: Large-Scale Video Panoptic Segmentation in the Wild: A Benchmark
http://openaccess.thecvf.com//content/CVPR2022/html/Miao_Large-Scale_Video_Panoptic_Segmentation_in_the_Wild_A_Benchmark_CVPR_2022_paper.html
AUTHORS: Jiaxu Miao, Xiaohan Wang, Yu Wu, Wei Li, Xu Zhang, Yunchao Wei, Yi Yang
HIGHLIGHT: In this paper, we present a new large-scale dataset for the video panoptic segmentation task, which aims to assign semantic classes and track identities to all pixels in a video.

1467, TITLE: Label, Verify, Correct: A Simple Few Shot Object Detection Method
http://openaccess.thecvf.com//content/CVPR2022/html/Kaul_Label_Verify_Correct_A_Simple_Few_Shot_Object_Detection_Method_CVPR_2022_paper.html
AUTHORS: Prannay Kaul, Weidi Xie, Andrew Zisserman
HIGHLIGHT: The objective of this paper is few-shot object detection (FSOD) - the task of expanding an object detector for a new category given only a few instances as training.

1468, TITLE: Aesthetic Text Logo Synthesis via Content-Aware Layout Inferring
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Aesthetic_Text_Logo_Synthesis_via_Content-Aware_Layout_Inferring_CVPR_2022_paper.html
AUTHORS: Yizhi Wang, Guo Pu, Wenhan Luo, Yexin Wang, Pengfei Xiong, Hongwen Kang, Zhouhui Lian
HIGHLIGHT: In this paper, we propose a content-aware layout generation network which takes glyph images and their corresponding text as input and synthesizes aesthetic layouts for them automatically.

1469, TITLE: Global Tracking via Ensemble of Local Trackers
http://openaccess.thecvf.com//content/CVPR2022/html/Zhou_Global_Tracking_via_Ensemble_of_Local_Trackers_CVPR_2022_paper.html
AUTHORS: Zikun Zhou, Jianqiu Chen, Wenjie Pei, Kaige Mao, Hongpeng Wang, Zhenyu He
HIGHLIGHT: In this work, we combine the advantages of both strategies: tracking the target in a global view while exploiting the temporal context.

1470, TITLE: Autoregressive Image Generation Using Residual Quantization
http://openaccess.thecvf.com//content/CVPR2022/html/Lee_Autoregressive_Image_Generation_Using_Residual_Quantization_CVPR_2022_paper.html
AUTHORS: Doyup Lee, Chiheon Kim, Saehoon Kim, Minsu Cho, Wook-Shin Han
HIGHLIGHT: In this study, we propose the two-stage framework, which consists of Residual-Quantized VAE (RQ-VAE) and RQ-Transformer, to effectively generate high-resolution images.

1471, TITLE: MPC: Multi-View Probabilistic Clustering
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_MPC_Multi-View_Probabilistic_Clustering_CVPR_2022_paper.html
AUTHORS: Junjie Liu, Junlong Liu, Shaotian Yan, Rongxin Jiang, Xiang Tian, Boxuan Gu, Yaowu Chen, Chen Shen, Jianqiang Huang
HIGHLIGHT: In this paper, we propose a novel unified framework for incomplete and complete MVC named multi-view probabilistic clustering (MPC).

1472, TITLE: End-to-End Compressed Video Representation Learning for Generic Event Boundary Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Li_End-to-End_Compressed_Video_Representation_Learning_for_Generic_Event_Boundary_Detection_CVPR_2022_paper.html
AUTHORS: Congcong Li, Xinyao Wang, Longyin Wen, Dexiang Hong, Tiejian Luo, Libo Zhang
HIGHLIGHT: To that end, we propose a new end-to-end compressed video representation learning for event boundary detection that leverages the rich information in the compressed domain, i.e., RGB, motion vectors, residuals, and the internal group of pictures (GOP) structure, without fully decoding the video.

1473, TITLE: GrainSpace: A Large-Scale Dataset for Fine-Grained and Domain-Adaptive Recognition of Cereal Grains
http://openaccess.thecvf.com//content/CVPR2022/html/Fan_GrainSpace_A_Large-Scale_Dataset_for_Fine-Grained_and_Domain-Adaptive_Recognition_of_CVPR_2022_paper.html
AUTHORS: Lei Fan, Yiwen Ding, Dongdong Fan, Donglin Di, Maurice Pagnucco, Yang Song
HIGHLIGHT: In this paper we formulate GAI as three ubiquitous computer vision tasks: fine-grained recognition, domain adaptation and out-of-distribution recognition.

1474, TITLE: BokehMe: When Neural Rendering Meets Classical Rendering

http://openaccess.thecvf.com//content/CVPR2022/html/Peng_BokehMe_When_Neural_Rendering_Meets_Classical_Rendering_CVPR_2022_paper.html
AUTHORS: Juewen Peng, Zhiguo Cao, Xianrui Luo, Hao Lu, Ke Xian, Jianming Zhang
HIGHLIGHT: We propose BokehMe, a hybrid bokeh rendering framework that marries a neural renderer with a classical physically motivated renderer.

1475, TITLE: Learning Modal-Invariant and Temporal-Memory for Video-Based Visible-Infrared Person Re-Identification
http://openaccess.thecvf.com//content/CVPR2022/html/Lin_Learning_Modal-Invariant_and_Temporal-Memory_for_Video-Based_Visible-Infrared_Person_Re-Identification_CVPR_2022_paper.html
AUTHORS: Xinyu Lin, Jinxing Li, Zeyu Ma, Huafeng Li, Shuang Li, Kaixiong Xu, Guangming Lu, David Zhang
HIGHLIGHT: In this paper, we primarily study the video-based cross-modal person Re-ID method.

1476, TITLE: MSDN: Mutually Semantic Distillation Network for Zero-Shot Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_MSDN_Mutually_Semantic_Distillation_Network_for_Zero-Shot_Learning_CVPR_2022_paper.html
AUTHORS: Shiming Chen, Ziming Hong, Guo-Sen Xie, Wenhan Yang, Qinmu Peng, Kai Wang, Jian Zhao, Xinge You
HIGHLIGHT: Prior works either simply align the global features of an image with its associated class semantic vector or utilize unidirectional attention to learn the limited latent semantic representations, which could not effectively discover the intrinsic semantic knowledge (e.g., attribute semantics) between visual and attribute features. To solve the above dilemma, we propose a Mutually Semantic Distillation Network (MSDN), which progressively distills the intrinsic semantic representations between visual and attribute features for ZSL.

1477, TITLE: Oriented RepPoints for Aerial Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Oriented_RepPoints_for_Aerial_Object_Detection_CVPR_2022_paper.html
AUTHORS: Wentong Li, Yijie Chen, Kaixuan Hu, Jianke Zhu
HIGHLIGHT: Unlike the mainstreamed approaches regressing the bounding box orientations, this paper proposes an effective adaptive points learning approach to aerial object detection by taking advantage of the adaptive points representation, which is able to capture the geometric information of the arbitrary-oriented instances.

1478, TITLE: OccAM's Laser: Occlusion-Based Attribution Maps for 3D Object Detectors on LiDAR Data
http://openaccess.thecvf.com//content/CVPR2022/html/Schinagl_OccAMs_Laser_Occlusion-Based_Attribution_Maps_for_3D_Object_Detectors_on_CVPR_2022_paper.html
AUTHORS: David Schinagl, Georg Krispel, Horst Possegger, Peter M. Roth, Horst Bischof
HIGHLIGHT: While 3D object detection in LiDAR point clouds is well-established in academia and industry, the explainability of these models is a largely unexplored field. In this paper, we propose a method to generate attribution maps for the detected objects in order to better understand the behavior of such models.

1479, TITLE: BigDatasetGAN: Synthesizing ImageNet With Pixel-Wise Annotations
http://openaccess.thecvf.com//content/CVPR2022/html/Li_BigDatasetGAN_Synthesizing_ImageNet_With_Pixel-Wise_Annotations_CVPR_2022_paper.html
AUTHORS: Daiqing Li, Huan Ling, Seung Wook Kim, Karsten Kreis, Sanja Fidler, Antonio Torralba
HIGHLIGHT: Here, we scale DatasetGAN to ImageNet scale of class diversity.

1480, TITLE: Align Representations With Base: A New Approach to Self-Supervised Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Align_Representations_With_Base_A_New_Approach_to_Self-Supervised_Learning_CVPR_2022_paper.html
AUTHORS: Shaofeng Zhang, Lyn Qiu, Feng Zhu, Junchi Yan, Hengrui Zhang, Rui Zhao, Hongyang Li, Xiaokang Yang
HIGHLIGHT: Specifically, the proposed intermediate variables are the nearest group of base vectors to representations. Hence, we call the proposed method ARB (Align Representations with Base).

1481, TITLE: Exploring Denoised Cross-Video Contrast for Weakly-Supervised Temporal Action Localization
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Exploring_Denoised_Cross-Video_Contrast_for_Weakly-Supervised_Temporal_Action_Localization_CVPR_2022_paper.html
AUTHORS: Jingjing Li, Tianyu Yang, Wei Ji, Jue Wang, Li Cheng
HIGHLIGHT: Inspired by recent success in unsupervised contrastive representation learning, we propose a novel denoised cross-video contrastive algorithm, aiming to enhance the feature discrimination ability of video snippets for accurate temporal action localization in the weakly-supervised setting.

1482, TITLE: Pre-Train, Self-Train, Distill: A Simple Recipe for Supersizing 3D Reconstruction
http://openaccess.thecvf.com//content/CVPR2022/html/Alwala_Pre-Train_Self-Train_Distill_A_Simple_Recipe_for_Supersizing_3D_Reconstruction_CVPR_2022_paper.html
AUTHORS: Kalyan Vasudev Alwala, Abhinav Gupta, Shubham Tulsiani
HIGHLIGHT: Our work learns a unified model for single-view 3D reconstruction of objects from hundreds of semantic categories.

1483, TITLE: Meta Distribution Alignment for Generalizable Person Re-Identification
http://openaccess.thecvf.com//content/CVPR2022/html/Ni_Meta_Distribution_Alignment_for_Generalizable_Person_Re-Identification_CVPR_2022_paper.html
AUTHORS: Hao Ni, Jingkuan Song, Xiaopeng Luo, Feng Zheng, Wen Li, Heng Tao Shen

HIGHLIGHT: However, they do not consider the target domain information which is unavailable in the training phrase of DG. To address this issue, we propose a novel Meta Distribution Alignment (MDA) method to enable them to share similar distribution in a test-time-training fashion.

1484, TITLE: TeachAugment: Data Augmentation Optimization Using Teacher Knowledge
http://openaccess.thecvf.com//content/CVPR2022/html/Suzuki_TeachAugment_Data_Augmentation_Optimization_Using_Teacher_Knowledge_CVPR_2022_paper.html
AUTHORS: Teppei Suzuki
HIGHLIGHT: In this paper, we propose a data augmentation optimization method based on the adversarial strategy called TeachAugment, which can produce informative transformed images to the model without requiring careful tuning by leveraging a teacher model.

1485, TITLE: SVIP: Sequence VerIfication for Procedures in Videos
http://openaccess.thecvf.com//content/CVPR2022/html/Qian_SVIP_Sequence_VerIfication_for_Procedures_in_Videos_CVPR_2022_paper.html
AUTHORS: Yicheng Qian, Weixin Luo, Dongze Lian, Xu Tang, Peilin Zhao, Shenghua Gao
HIGHLIGHT: In this paper, we propose a novel sequence verification task that aims to distinguish positive video pairs performing the same action sequence from negative ones with step-level transformations but still conducting the same task.

1486, TITLE: Weakly Supervised Temporal Sentence Grounding With Gaussian-Based Contrastive Proposal Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Zheng_Weakly_Supervised_Temporal_Sentence_Grounding_With_Gaussian-Based_Contrastive_Proposal_Learning_CVPR_2022_paper.html
AUTHORS: Minghang Zheng, Yanjie Huang, Qingchao Chen, Yuxin Peng, Yang Liu
HIGHLIGHT: Moreover, they train their model to distinguish positive visual-language pairs from negative ones randomly collected from other videos, ignoring the highly confusing video segments within the same video. In this paper, we propose Contrastive Proposal Learning (CPL) to overcome the above limitations.

1487, TITLE: Low-Resource Adaptation for Personalized Co-Speech Gesture Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Ahuja_Low-Resource_Adaptation_for_Personalized_Co-Speech_Gesture_Generation_CVPR_2022_paper.html
AUTHORS: Chaitanya Ahuja, Dong Won Lee, Louis-Philippe Morency
HIGHLIGHT: We propose an approach, named DiffGAN, that efficiently personalizes co-speech gesture generation models of a high-resource source speaker to target speaker with just 2 minutes of target training data.

1488, TITLE: BoosterNet: Improving Domain Generalization of Deep Neural Nets Using Culpability-Ranked Features
http://openaccess.thecvf.com//content/CVPR2022/html/Bayasi_BoosterNet_Improving_Domain_Generalization_of_Deep_Neural_Nets_Using_Culpability-Ranked_CVPR_2022_paper.html
AUTHORS: Nourhan Bayasi, Ghassan Hamarneh, Rafeef Garbi
HIGHLIGHT: In this work, we propose BoosterNet, a lean add-on network that can be simply appended to any arbitrary core network to improve its generalization capability without requiring any changes in its architecture or training procedure.

1489, TITLE: Task-Specific Inconsistency Alignment for Domain Adaptive Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Zhao_Task-Specific_Inconsistency_Alignment_for_Domain_Adaptive_Object_Detection_CVPR_2022_paper.html
AUTHORS: Liang Zhao, Limin Wang
HIGHLIGHT: To alleviate this problem of domain shift, conventional wisdom typically concentrates solely on reducing the discrepancy between the source and target domains via attached domain classifiers, yet ignoring the difficulty of such transferable features in coping with both classification and localization subtasks in object detection. To address this issue, in this paper, we propose Task-specific Inconsistency Alignment (TIA), by developing a new alignment mechanism in separate task spaces, improving the performance of the detector on both subtasks.

1490, TITLE: HDR-NeRF: High Dynamic Range Neural Radiance Fields
http://openaccess.thecvf.com//content/CVPR2022/html/Huang_HDR-NeRF_High_Dynamic_Range_Neural_Radiance_Fields_CVPR_2022_paper.html
AUTHORS: Xin Huang, Qi Zhang, Ying Feng, Hongdong Li, Xuan Wang, Qing Wang
HIGHLIGHT: We present High Dynamic Range Neural Radiance Fields (HDR-NeRF) to recover an HDR radiance field from a set of low dynamic range (LDR) views with different exposures.

1491, TITLE: MS2DG-Net: Progressive Correspondence Learning via Multiple Sparse Semantics Dynamic Graph
http://openaccess.thecvf.com//content/CVPR2022/html/Dai_MS2DG-Net_Progressive_Correspondence_Learning_via_Multiple_Sparse_Semantics_Dynamic_Graph_CVPR_2022_paper.html
AUTHORS: Luanyuan Dai, Yizhang Liu, Jiayi Ma, Lifang Wei, Taotao Lai, Changcai Yang, Riqing Chen
HIGHLIGHT: However, most such works ignore similar sparse semantics information between two given images and cannot capture local topology among correspondences well. Therefore, to deal with the above problems, Multiple Sparse Semantics Dynamic Graph Network (MS^ 2 DG-Net) is proposed, in this paper, to predict probabilities of correspondences as inliers and recover camera poses.

1492, TITLE: Neural Emotion Director: Speech-Preserving Semantic Control of Facial Expressions in "In-the-Wild" Videos

http://openaccess.thecvf.com//content/CVPR2022/html/Papantoniou_Neural_Emotion_Director_Speech-Preserving_Semantic_Control_of_Facial_Expressions_in_CVPR_2022_paper.html
AUTHORS: Foivos Paraperas Papantoniou, Panagiotis P. Filntisis, Petros Maragos, Anastasios Roussos
HIGHLIGHT: In this paper, we introduce a novel deep learning method for photo-realistic manipulation of the emotional state of actors in "in-the-wild" videos.

1493, TITLE: Learning To Listen: Modeling Non-Deterministic Dyadic Facial Motion
http://openaccess.thecvf.com//content/CVPR2022/html/Ng_Learning_To_Listen_Modeling_Non-Deterministic_Dyadic_Facial_Motion_CVPR_2022_paper.html
AUTHORS: Evonne Ng, Hanbyul Joo, Liwen Hu, Hao Li, Trevor Darrell, Angjoo Kanazawa, Shiry Ginosar
HIGHLIGHT: We present a framework for modeling interactional communication in dyadic conversations: given multimodal inputs of a speaker, we autoregressively output multiple possibilities of corresponding listener motion.

1494, TITLE: 3PSDF: Three-Pole Signed Distance Function for Learning Surfaces With Arbitrary Topologies
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_3PSDF_Three-Pole_Signed_Distance_Function_for_Learning_Surfaces_With_Arbitrary_CVPR_2022_paper.html
AUTHORS: Weikai Chen, Cheng Lin, Weiyang Li, Bo Yang
HIGHLIGHT: We present a novel learnable implicit representation, called the three-pole signed distance function (3PSDF), that can represent non-watertight 3D shapes with arbitrary topologies while supporting easy field-to-mesh conversion using the classic Marching Cubes algorithm.

1495, TITLE: Capturing Humans in Motion: Temporal-Attentive 3D Human Pose and Shape Estimation From Monocular Video
http://openaccess.thecvf.com//content/CVPR2022/html/Wei_Capturing_Humans_in_Motion_Temporal-Attentive_3D_Human_Pose_and_Shape_CVPR_2022_paper.html
AUTHORS: Wen-Li Wei, Jen-Chun Lin, Tyng-Luh Liu, Hong-Yuan Mark Liao
HIGHLIGHT: However, the existing methods mainly rely on recurrent or convolutional operation to model such temporal information, which limits the ability to capture non-local context relations of human motion. To address this problem, we propose a motion pose and shape network (MPS-Net) to effectively capture humans in motion to estimate accurate and temporally coherent 3D human pose and shape from a video.

1496, TITLE: MixFormer: End-to-End Tracking With Iterative Mixed Attention
http://openaccess.thecvf.com//content/CVPR2022/html/Cui_MixFormer_End-to-End_Tracking_With_Iterative_Mixed_Attention_CVPR_2022_paper.html
AUTHORS: Yutao Cui, Cheng Jiang, Limin Wang, Gangshan Wu
HIGHLIGHT: To simplify this pipeline and unify the process of feature extraction and target information integration, we present a compact tracking framework, termed as MixFormer, built upon transformers.

1497, TITLE: Sparse Fuse Dense: Towards High Quality 3D Detection With Depth Completion
http://openaccess.thecvf.com//content/CVPR2022/html/Wu_Sparse_Fuse_Dense_Towards_High_Quality_3D_Detection_With_Depth_CVPR_2022_paper.html
AUTHORS: Xiaopei Wu, Liang Peng, Honghui Yang, Liang Xie, Chenxi Huang, Chengqi Deng, Haifeng Liu, Deng Cai
HIGHLIGHT: In this paper, we present a novel multi-modal framework SFD (Sparse Fuse Dense), which utilizes pseudo point clouds generated from depth completion to tackle the issues mentioned above.

1498, TITLE: GIRAFFE HD: A High-Resolution 3D-Aware Generative Model
http://openaccess.thecvf.com//content/CVPR2022/html/Xue_GIRAFFE_HD_A_High-Resolution_3D-Aware_Generative_Model_CVPR_2022_paper.html
AUTHORS: Yang Xue, Yuheng Li, Krishna Kumar Singh, Yong Jae Lee
HIGHLIGHT: We propose GIRAFFE HD, a high-resolution 3D-aware generative model that inherits all of GIRAFFE's controllable features while generating high-quality, high-resolution images (512^2 resolution and above).

1499, TITLE: InOut: Diverse Image Outpainting via GAN Inversion
http://openaccess.thecvf.com//content/CVPR2022/html/Cheng_InOut_Diverse_Image_Outpainting_via_GAN_Inversion_CVPR_2022_paper.html
AUTHORS: Yen-Chi Cheng, Chieh Hubert Lin, Hsin-Ying Lee, Jian Ren, Sergey Tulyakov, Ming-Hsuan Yang
HIGHLIGHT: In this work, we formulate the problem from the perspective of inverting generative adversarial networks.

1500, TITLE: PNP: Robust Learning From Noisy Labels by Probabilistic Noise Prediction
http://openaccess.thecvf.com//content/CVPR2022/html/Sun_PNP_Robust_Learning_From_Noisy_Labels_by_Probabilistic_Noise_Prediction_CVPR_2022_paper.html
AUTHORS: Zeren Sun, Fumin Shen, Dan Huang, Qiong Wang, Xiangbo Shu, Yazhou Yao, Jinhui Tang
HIGHLIGHT: However, these approaches focus on dividing samples by order sorting or threshold selection, inevitably introducing hyper-parameters (e.g., selection ratio / threshold) that are hard-to-tune and dataset-dependent. To this end, we propose a simple yet effective approach named PNP (Probabilistic Noise Prediction) to explicitly model label noise.

1501, TITLE: Estimating Structural Disparities for Face Models
http://openaccess.thecvf.com//content/CVPR2022/html/Ardeshir_Estimating_Structural_Disparities_for_Face_Models_CVPR_2022_paper.html
AUTHORS: Shervin Ardeshir, Cristina Segalin, Nathan Kallus

HIGHLIGHT: On the other hand, in many scenarios noisy groupings may be obtainable using some form of a proxy, which would allow measuring disparity metrics across sub-populations. Here we explore performing such analysis on computer vision models trained on human faces, and on tasks such as face attribute prediction and affect estimation.

1502, TITLE: Revisiting the Transferability of Supervised Pretraining: An MLP Perspective
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Revisiting_the_Transferability_of_Supervised_Pretraining_An_MLP_Perspective_CVPR_2022_paper.html
AUTHORS: Yizhou Wang, Shixiang Tang, Feng Zhu, Lei Bai, Rui Zhao, Donglian Qi, Wanli Ouyang
HIGHLIGHT: Recent progress on unsupervised pretraining methods shows superior transfer performance to their supervised counterparts. This paper revisits this phenomenon and sheds new light on understanding the transferability gap between unsupervised and supervised pretraining from a multilayer perceptron (MLP) perspective.

1503, TITLE: Plenoxels: Radiance Fields Without Neural Networks
http://openaccess.thecvf.com//content/CVPR2022/html/Fridovich-Keil_Plenoxels_Radiance_Fields_Without_Neural_Networks_CVPR_2022_paper.html
AUTHORS: Sara Fridovich-Keil, Alex Yu, Matthew Tancik, Qinhong Chen, Benjamin Recht, Angjoo Kanazawa
HIGHLIGHT: We introduce Plenoxels (plenoptic voxels), a system for photorealistic view synthesis.

1504, TITLE: What Matters for Meta-Learning Vision Regression Tasks?
http://openaccess.thecvf.com//content/CVPR2022/html/Gao_What_Matters_for_Meta-Learning_Vision_Regression_Tasks_CVPR_2022_paper.html
AUTHORS: Ning Gao, Hanna Ziesche, Ngo Anh Vien, Michael Volpp, Gerhard Neumann
HIGHLIGHT: Meta-learning is widely used in few-shot classification and function regression due to its ability to quickly adapt to unseen tasks. However, it has not yet been well explored on regression tasks with high dimensional inputs such as images. This paper makes two main contributions that help understand this barely explored area.

1505, TITLE: Knowledge-Driven Self-Supervised Representation Learning for Facial Action Unit Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Chang_Knowledge-Driven_Self-Supervised_Representation_Learning_for_Facial_Action_Unit_Recognition_CVPR_2022_paper.html
AUTHORS: Yanan Chang, Shangfei Wang
HIGHLIGHT: However, the complex labeling process makes it challenging to provide AU annotations for large amounts of facial images. To remedy this, we utilize AU labeling rules defined by the Facial Action Coding System (FACS) to design a novel knowledge-driven self-supervised representation learning framework for AU recognition.

1506, TITLE: Selective-Supervised Contrastive Learning With Noisy Labels
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Selective-Supervised_Contrastive_Learning_With_Noisy_Labels_CVPR_2022_paper.html
AUTHORS: Shikun Li, Xiaobo Xia, Shiming Ge, Tongliang Liu
HIGHLIGHT: To learn robust representations and handle noisy labels, we propose selective-supervised contrastive learning (Sel-CL) in this paper.

1507, TITLE: Learning Second Order Local Anomaly for General Face Forgery Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Fei_Learning_Second_Order_Local_Anomaly_for_General_Face_Forgery_Detection_CVPR_2022_paper.html
AUTHORS: Jianwei Fei, Yunshu Dai, Peipeng Yu, Tianrun Shen, Zhihua Xia, Jian Weng
HIGHLIGHT: In this work, we propose a novel method to improve the generalization ability of CNN-based face forgery detectors.

1508, TITLE: ADAS: A Direct Adaptation Strategy for Multi-Target Domain Adaptive Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Lee_ADAS_A_Direct_Adaptation_Strategy_for_Multi-Target_Domain_Adaptive_Semantic_CVPR_2022_paper.html
AUTHORS: Seunghun Lee, Wonhyeok Choi, Changjae Kim, Minwoo Choi, Sunghoon Im
HIGHLIGHT: In this paper, we present a direct adaptation strategy (ADAS), which aims to directly adapt a single model to multiple target domains in a semantic segmentation task without pretrained domain-specific models.

1509, TITLE: The Devil Is in the Labels: Noisy Label Correction for Robust Scene Graph Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Li_The_Devil_Is_in_the_Labels_Noisy_Label_Correction_for_CVPR_2022_paper.html
AUTHORS: Lin Li, Long Chen, Yifeng Huang, Zhimeng Zhang, Songyang Zhang, Jun Xiao
HIGHLIGHT: However, almost all existing SGG models have overlooked the ground-truth annotation qualities of prevailing SGG datasets, i.e., they always assume: 1) all the manually annotated positive samples are equally correct; 2) all the un-annotated negative samples are absolutely background. In this paper, we argue that both assumptions are inapplicable to SGG: there are numerous "noisy" ground-truth predicate labels that break these two assumptions, and these noisy samples actually harm the training of unbiased SGG models.

1510, TITLE: LAVT: Language-Aware Vision Transformer for Referring Image Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_LAVT_Language-Aware_Vision_Transformer_for_Referring_Image_Segmentation_CVPR_2022_paper.html
AUTHORS: Zhao Yang, Jiaqi Wang, Yansong Tang, Kai Chen, Hengshuang Zhao, Philip H.S. Torr

HIGHLIGHT: Adopting a different approach in this work, we show that significantly better cross-modal alignments can be achieved through the early fusion of linguistic and visual features in intermediate layers of a vision Transformer encoder network.

1511, TITLE: SimT: Handling Open-Set Noise for Domain Adaptive Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Guo_SimT_Handling_Open-Set_Noise_for_Domain_Adaptive_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Xiaoqing Guo, Jie Liu, Tongliang Liu, Yixuan Yuan
HIGHLIGHT: In this paper, we propose a simplex noise transition matrix (SimT) to model the mixed noise distributions in DA semantic segmentation and formulate the problem as estimation of SimT.

1512, TITLE: Interspace Pruning: Using Adaptive Filter Representations To Improve Training of Sparse CNNs
http://openaccess.thecvf.com//content/CVPR2022/html/Wimmer_Interspace_Pruning_Using_Adaptive_Filter_Representations_To_Improve_Training_of_CVPR_2022_paper.html
AUTHORS: Paul Wimmer, Jens Mehnert, Alexandru Condurache
HIGHLIGHT: In this work, we provide mathematical evidence for IP's superior performance and demonstrate that IP outperforms SP on all tested state-of-the-art unstructured pruning methods.

1513, TITLE: PLAD: Learning To Infer Shape Programs With Pseudo-Labels and Approximate Distributions
http://openaccess.thecvf.com//content/CVPR2022/html/Jones_PLAD_Learning_To_Infer_Shape_Programs_With_Pseudo-Labels_and_Approximate_CVPR_2022_paper.html
AUTHORS: R. Kenny Jones, Homer Walke, Daniel Ritchie
HIGHLIGHT: In self-training, shapes are passed through a recognition model, which predicts programs that are treated as 'pseudo-labels' for those shapes. Related to these approaches, we introduce a novel self-training variant unique to program inference, where program pseudo-labels are paired with their executed output shapes, avoiding label mismatch at the cost of an approximate shape distribution.

1514, TITLE: PTTR: Relational 3D Point Cloud Object Tracking With Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Zhou_PTTR_Relational_3D_Point_Cloud_Object_Tracking_With_Transformer_CVPR_2022_paper.html
AUTHORS: Changqing Zhou, Zhipeng Luo, Yueru Luo, Tianrui Liu, Liang Pan, Zhongang Cai, Haiyu Zhao, Shijian Lu
HIGHLIGHT: Motivated by the success of transformers, we propose Point Tracking TRansformer (PTTR), which efficiently predicts high-quality 3D tracking results in a coarse-to-fine manner with the help of transformer operations.

1515, TITLE: Frequency-Driven Imperceptible Adversarial Attack on Semantic Similarity
http://openaccess.thecvf.com//content/CVPR2022/html/Luo_Frequency-Driven_Imperceptible_Adversarial_Attack_on_Semantic_Similarity_CVPR_2022_paper.html
AUTHORS: Cheng Luo, Qinliang Lin, Weicheng Xie, Bizhu Wu, Jinheng Xie, Linlin Shen
HIGHLIGHT: To circumvent the former problem, we propose a novel algorithm that attacks semantic similarity on feature representations.

1516, TITLE: ZZ-Net: A Universal Rotation Equivariant Architecture for 2D Point Clouds
http://openaccess.thecvf.com//content/CVPR2022/html/Bokman_ZZ-Net_A_Universal_Rotation_Equivariant_Architecture_for_2D_Point_Clouds_CVPR_2022_paper.html
AUTHORS: Georg Bökman, Fredrik Kahl, Axel Flinth
HIGHLIGHT: In this paper, we are concerned with rotation equivariance on 2D point cloud data.

1517, TITLE: Video Demoireing With Relation-Based Temporal Consistency
http://openaccess.thecvf.com//content/CVPR2022/html/Dai_Video_Demoireing_With_Relation-Based_Temporal_Consistency_CVPR_2022_paper.html
AUTHORS: Peng Dai, Xin Yu, Lan Ma, Baoheng Zhang, Jia Li, Wenbo Li, Jiajun Shen, Xiaojuan Qi
HIGHLIGHT: Considering the increasing demands for capturing videos, we study how to remove such undesirable moire patterns in videos, namely video demoireing. To this end, we introduce the first hand-held video demoireing dataset with a dedicated data collection pipeline to ensure spatial and temporal alignments of captured data.

1518, TITLE: Co-Domain Symmetry for Complex-Valued Deep Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Singhal_Co-Domain_Symmetry_for_Complex-Valued_Deep_Learning_CVPR_2022_paper.html
AUTHORS: Utkarsh Singhal, Yifei Xing, Stella X. Yu
HIGHLIGHT: We study complex-valued scaling as a type of symmetry natural and unique to complex-valued measurements and representations.

1519, TITLE: Industrial Style Transfer With Large-Scale Geometric Warping and Content Preservation
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Industrial_Style_Transfer_With_Large-Scale_Geometric_Warping_and_Content_Preservation_CVPR_2022_paper.html
AUTHORS: Jinchao Yang, Fei Guo, Shuo Chen, Jun Li, Jian Yang
HIGHLIGHT: We propose a novel style transfer method to quickly create a new visual product with a nice appearance for industrial designers' reference.

1520, TITLE: Modeling Image Composition for Complex Scene Generation

http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Modeling_Image_Composition_for_Complex_Scene_Generation_CVPR_2022_paper.html
AUTHORS: Zuopeng Yang, Daqing Liu, Chaoyue Wang, Jie Yang, Dacheng Tao
HIGHLIGHT: We present a method that achieves state-of-the-art results on challenging (few-shot) layout-to-image generation tasks by accurately modeling textures, structures and relationships contained in a complex scene.

1521, TITLE: SS3D: Sparsely-Supervised 3D Object Detection From Point Cloud
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_SS3D_Sparsely-Supervised_3D_Object_Detection_From_Point_Cloud_CVPR_2022_paper.html
AUTHORS: Chuandong Liu, Chenqiang Gao, Fangcen Liu, Jiang Liu, Deyu Meng, Xinbo Gao
HIGHLIGHT: In this paper, we propose a sparselysupervised 3D object detection method, named SS3D.

1522, TITLE: Remember the Difference: Cross-Domain Few-Shot Semantic Segmentation via Meta-Memory Transfer
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Remember_the_Difference_Cross-Domain_Few-Shot_Semantic_Segmentation_via_Meta-Memory_Transfer_CVPR_2022_paper.html
AUTHORS: Wenjian Wang, Lijuan Duan, Yuxi Wang, Qing En, Junsong Fan, Zhaoxiang Zhang
HIGHLIGHT: The actual domain shift problem significantly reduces the performance of few-shot learning. To remedy this problem, we propose an interesting and challenging cross-domain few-shot semantic segmentation task, where the training and test tasks perform on different domains.

1523, TITLE: GRAM: Generative Radiance Manifolds for 3D-Aware Image Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Deng_GRAM_Generative_Radiance_Manifolds_for_3D-Aware_Image_Generation_CVPR_2022_paper.html
AUTHORS: Yu Deng, Jiaolong Yang, Jianfeng Xiang, Xin Tong
HIGHLIGHT: We propose a novel approach that regulates point sampling and radiance field learning on 2D manifolds, embodied as a set of learned implicit surfaces in the 3D volume.

1524, TITLE: UniVIP: A Unified Framework for Self-Supervised Visual Pre-Training
http://openaccess.thecvf.com//content/CVPR2022/html/Li_UniVIP_A_Unified_Framework_for_Self-Supervised_Visual_Pre-Training_CVPR_2022_paper.html
AUTHORS: Zhaowen Li, Yousong Zhu, Fan Yang, Wei Li, Chaoyang Zhao, Yingying Chen, Zhiyang Chen, Jiahao Xie, Liwei Wu, Rui Zhao, Ming Tang, Jinqiao Wang
HIGHLIGHT: However, the success of popular SSL methods has limited on single-centric-object images like those in ImageNet and ignores the correlation among the scene and instances, as well as the semantic difference of instances in the scene. To address the above problems, we propose a Unified Self-supervised Visual Pre-training (UniVIP), a novel self-supervised framework to learn versatile visual representations on either single-centric-object or non-iconic dataset.

1525, TITLE: GraFormer: Graph-Oriented Transformer for 3D Pose Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Zhao_GraFormer_Graph-Oriented_Transformer_for_3D_Pose_Estimation_CVPR_2022_paper.html
AUTHORS: Weixi Zhao, Weiqiang Wang, Yunjie Tian
HIGHLIGHT: To better model the relation of joints for 3D pose estimation, we propose an effective but simple network, called GraFormer, where a novel transformer architecture is designed via embedding graph convolution layers after multi-head attention block.

1526, TITLE: Decoupling Zero-Shot Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Ding_Decoupling_Zero-Shot_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Jian Ding, Nan Xue, Gui-Song Xia, Dengxin Dai
HIGHLIGHT: Based on the decoupling formulation, we propose a simple and effective zero-shot semantic segmentation model, called ZegFormer, which outperforms the previous methods on ZS3 standard benchmarks by large margins, e.g., 22 points on the PAS-CAL VOC and 3 points on the COCO-Stuff in terms of mIoU for unseen classes.

1527, TITLE: Neural Collaborative Graph Machines for Table Structure Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Neural_Collaborative_Graph_Machines_for_Table_Structure_Recognition_CVPR_2022_paper.html
AUTHORS: Hao Liu, Xin Li, Bing Liu, Deqiang Jiang, Yinsong Liu, Bo Ren
HIGHLIGHT: With the aim of filling this gap, we present a novel Neural Collaborative Graph Machines (NCGM) equipped with stacked collaborative blocks, which alternatively extracts intra-modality context and models inter-modality interactions in a hierarchical way.

1528, TITLE: Towards Robust Vision Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Mao_Towards_Robust_Vision_Transformer_CVPR_2022_paper.html
AUTHORS: Xiaofeng Mao, Gege Qi, Yuefeng Chen, Xiaodan Li, Ranjie Duan, Shaokai Ye, Yuan He, Hui Xue
HIGHLIGHT: In this work, we conduct systematic evaluation on components of ViTs in terms of their impact on robustness to adversarial examples, common corruptions and distribution shifts.

1529, TITLE: DeepCurrents: Learning Implicit Representations of Shapes With Boundaries
http://openaccess.thecvf.com//content/CVPR2022/html/Palmer_DeepCurrents_Learning_Implicit_Representations_of_Shapes_With_Boundaries_CVPR_2022_paper.html

AUTHORS: David Palmer, Dmitriy Smirnov, Stephanie Wang, Albert Chern, Justin Solomon
HIGHLIGHT: We propose a hybrid shape representation that combines explicit boundary curves with implicit learned interiors.

1530, TITLE: Learning Affordance Grounding From Exocentric Images
http://openaccess.thecvf.com//content/CVPR2022/html/Luo_Learning_Affordance_Grounding_From_Exocentric_Images_CVPR_2022_paper.html
AUTHORS: Hongchen Luo, Wei Zhai, Jing Zhang, Yang Cao, Dacheng Tao
HIGHLIGHT: Human has the ability that transform the various exocentric interactions to invariant egocentric affordance so as to counter the impact of interactive diversity. To empower an agent with such ability, this paper proposes a task of affordance grounding from exocentric view, i.e., given exocentric human-object interaction and egocentric object images, learning the affordance knowledge of the object and transferring it to the egocentric image using only the affordance label as supervision.

1531, TITLE: Templates for 3D Object Pose Estimation Revisited: Generalization to New Objects and Robustness to Occlusions
http://openaccess.thecvf.com//content/CVPR2022/html/Nguyen_Templates_for_3D_Object_Pose_Estimation_Revisited_Generalization_to_New_CVPR_2022_paper.html
AUTHORS: Van Nguyen Nguyen, Yinlin Hu, Yang Xiao, Mathieu Salzmann, Vincent Lepetit
HIGHLIGHT: We present a method that can recognize new objects and estimate their 3D pose in RGB images even under partial occlusions.

1532, TITLE: Stochastic Variance Reduced Ensemble Adversarial Attack for Boosting the Adversarial Transferability
http://openaccess.thecvf.com//content/CVPR2022/html/Xiong_Stochastic_Variance_Reduced_Ensemble_Adversarial_Attack_for_Boosting_the_Adversarial_CVPR_2022_paper.html
AUTHORS: Yifeng Xiong, Jiadong Lin, Min Zhang, John E. Hopcroft, Kun He
HIGHLIGHT: In this work, we treat the iterative ensemble attack as a stochastic gradient descent optimization process, in which the variance of the gradients on different models may lead to poor local optima.

1533, TITLE: Unknown-Aware Object Detection: Learning What You Don't Know From Videos in the Wild
http://openaccess.thecvf.com//content/CVPR2022/html/Du_Unknown-Aware_Object_Detection_Learning_What_You_Dont_Know_From_Videos_CVPR_2022_paper.html
AUTHORS: Xuefeng Du, Xin Wang, Gabriel Gozum, Yixuan Li
HIGHLIGHT: We propose a new unknown-aware object detection framework through Spatial-Temporal Unknown Distillation (STUD), which distills unknown objects from videos in the wild and meaningfully regularizes the model's decision boundary.

1534, TITLE: Multi-Modal Extreme Classification
http://openaccess.thecvf.com//content/CVPR2022/html/Mittal_Multi-Modal_Extreme_Classification_CVPR_2022_paper.html
AUTHORS: Anshul Mittal, Kunal Dahiya, Shreya Malani, Janani Ramaswamy, Seba Kuruvilla, Jitendra Ajmera, Keng-hao Chang, Sumeet Agarwal, Purushottam Kar, Manik Varma
HIGHLIGHT: This paper develops the MUFIN technique for extreme classification (XC) tasks with millions of labels where datapoints and labels are endowed with visual and textual descriptors.

1535, TITLE: IFOR: Iterative Flow Minimization for Robotic Object Rearrangement
http://openaccess.thecvf.com//content/CVPR2022/html/Goyal_IFOR_Iterative_Flow_Minimization_for_Robotic_Object_Rearrangement_CVPR_2022_paper.html
AUTHORS: Ankit Goyal, Arsalan Mousavian, Chris Paxton, Yu-Wei Chao, Brian Okorn, Jia Deng, Dieter Fox
HIGHLIGHT: We propose IFOR, Iterative Flow Minimization for Robotic Object Rearrangement, an end-to-end method for the challenging problem of object rearrangement for unknown objects given an RGBD image of the original and final scenes.

1536, TITLE: Training-Free Transformer Architecture Search
http://openaccess.thecvf.com//content/CVPR2022/html/Zhou_Training-Free_Transformer_Architecture_Search_CVPR_2022_paper.html
AUTHORS: Qinqin Zhou, Kekai Sheng, Xiawu Zheng, Ke Li, Xing Sun, Yonghong Tian, Jie Chen, Rongrong Ji
HIGHLIGHT: In this paper, for the first time, we investigate how to conduct TAS in a training-free manner and devise an effective training-free TAS (TF-TAS) scheme.

1537, TITLE: Zero Experience Required: Plug & Play Modular Transfer Learning for Semantic Visual Navigation
http://openaccess.thecvf.com//content/CVPR2022/html/Al-Halah_Zero_Experience_Required_Plug__Play_Modular_Transfer_Learning_for_CVPR_2022_paper.html
AUTHORS: Ziad Al-Halah, Santhosh Kumar Ramakrishnan, Kristen Grauman
HIGHLIGHT: We present a unified approach to visual navigation using a novel modular transfer learning model.

1538, TITLE: Non-Isotropy Regularization for Proxy-Based Deep Metric Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Roth_Non-Isotropy_Regularization_for_Proxy-Based_Deep_Metric_Learning_CVPR_2022_paper.html
AUTHORS: Karsten Roth, Oriol Vinyals, Zeynep Akata
HIGHLIGHT: Given the inherent non-bijectiveness of used distance functions, this can induce locally isotropic sample distributions, leading to crucial semantic context being missed due to difficulties resolving local structures and intraclass relations between samples. To alleviate this problem, we propose non-isotropy regularization (NIR) for proxy-based Deep Metric Learning.

1539, TITLE: C2AM: Contrastive Learning of Class-Agnostic Activation Map for Weakly Supervised Object Localization and Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Xie_C2AM_Contrastive_Learning_of_Class-Agnostic_Activation_Map_for_Weakly_Supervised_CVPR_2022_paper.html
AUTHORS: Jinheng Xie, Jianfeng Xiang, Junliang Chen, Xianxu Hou, Xiaodong Zhao, Linlin Shen
HIGHLIGHT: In this paper, we propose Contrastive learning for Class-agnostic Activation Map (C^2AM) generation only using unlabeled image data, without the involvement of image-level supervision.

1540, TITLE: TopFormer: Token Pyramid Transformer for Mobile Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_TopFormer_Token_Pyramid_Transformer_for_Mobile_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Wenqiang Zhang, Zilong Huang, Guozhong Luo, Tao Chen, Xinggang Wang, Wenyu Liu, Gang Yu, Chunhua Shen
HIGHLIGHT: In this paper, we present a mobile-friendly architecture named Token Pyramid Vision Transformer (TopFormer).

1541, TITLE: 3DAC: Learning Attribute Compression for Point Clouds
http://openaccess.thecvf.com//content/CVPR2022/html/Fang_3DAC_Learning_Attribute_Compression_for_Point_Clouds_CVPR_2022_paper.html
AUTHORS: Guangchi Fang, Qingyong Hu, Hanyun Wang, Yiling Xu, Yulan Guo
HIGHLIGHT: Through an in-depth exploration of the relationships between different encoding steps and different attribute channels, we introduce a deep compression network, termed 3DAC, to explicitly compress the attributes of 3D point clouds and reduce storage usage in this paper.

1542, TITLE: Learning a Structured Latent Space for Unsupervised Point Cloud Completion
http://openaccess.thecvf.com//content/CVPR2022/html/Cai_Learning_a_Structured_Latent_Space_for_Unsupervised_Point_Cloud_Completion_CVPR_2022_paper.html
AUTHORS: Yingjie Cai, Kwan-Yee Lin, Chao Zhang, Qiang Wang, Xiaogang Wang, Hongsheng Li
HIGHLIGHT: In this work, we propose a novel framework, which learns a unified and structured latent space that encoding both partial and complete point clouds.

1543, TITLE: The Wanderings of Odysseus in 3D Scenes
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_The_Wanderings_of_Odysseus_in_3D_Scenes_CVPR_2022_paper.html
AUTHORS: Yan Zhang, Siyu Tang
HIGHLIGHT: Our goal is to populate digital environments, in which digital humans have diverse body shapes, move perpetually, and have plausible body-scene contact.

1544, TITLE: Few-Shot Learning With Noisy Labels
http://openaccess.thecvf.com//content/CVPR2022/html/Liang_Few-Shot_Learning_With_Noisy_Labels_CVPR_2022_paper.html
AUTHORS: Kevin J. Liang, Samrudhdhi B. Rangrej, Vladan Petrovic, Tal Hassner
HIGHLIGHT: To address mislabeled samples in FSL settings, we make several technical contributions.

1545, TITLE: Understanding 3D Object Articulation in Internet Videos
http://openaccess.thecvf.com//content/CVPR2022/html/Qian_Understanding_3D_Object_Articulation_in_Internet_Videos_CVPR_2022_paper.html
AUTHORS: Shengyi Qian, Linyi Jin, Chris Rockwell, Siyi Chen, David F. Fouhey
HIGHLIGHT: We propose to investigate detecting and characterizing the 3D planar articulation of objects from ordinary RGB videos.

1546, TITLE: Multi-Level Representation Learning With Semantic Alignment for Referring Video Object Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Wu_Multi-Level_Representation_Learning_With_Semantic_Alignment_for_Referring_Video_Object_CVPR_2022_paper.html
AUTHORS: Dongming Wu, Xingping Dong, Ling Shao, Jianbing Shen
HIGHLIGHT: The limitation of visual representation is prone to causing vision-language mismatching and producing poor segmentation results. To address this, we propose a novel multi-level representation learning approach, which explores the inherent structure of the video content to provide a set of discriminative visual embedding, enabling more effective vision-language semantic alignment.

1547, TITLE: Paramixer: Parameterizing Mixing Links in Sparse Factors Works Better Than Dot-Product Self-Attention
http://openaccess.thecvf.com//content/CVPR2022/html/Yu_Paramixer_Parameterizing_Mixing_Links_in_Sparse_Factors_Works_Better_Than_CVPR_2022_paper.html
AUTHORS: Tong Yu, Ruslan Khalitov, Lei Cheng, Zhirong Yang
HIGHLIGHT: In this paper, we propose a novel scalable and effective mixing building block called Paramixer.

1548, TITLE: Interactive Image Synthesis With Panoptic Layout Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Interactive_Image_Synthesis_With_Panoptic_Layout_Generation_CVPR_2022_paper.html
AUTHORS: Bo Wang, Tao Wu, Minfeng Zhu, Peng Du

HIGHLIGHT: When placement of bounding boxes is subject to perturbation, layout-based models suffer from "missing regions" in the constructed semantic layouts and hence undesirable artifacts in the generated images. In this work, we propose Panoptic Layout Generative Adversarial Network (PLGAN) to address this challenge.

1549, TITLE: Pseudo-Stereo for Monocular 3D Object Detection in Autonomous Driving
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Pseudo-Stereo_for_Monocular_3D_Object_Detection_in_Autonomous_Driving_CVPR_2022_paper.html
AUTHORS: Yi-Nan Chen, Hang Dai, Yong Ding
HIGHLIGHT: The gap in image-to-image generation for stereo views is much smaller than that in image-to-LiDAR generation. Motivated by this, we propose a Pseudo-Stereo 3D detection framework with three novel virtual view generation methods, including image-level generation, feature-level generation, and feature-clone, for detecting 3D objects from a single image.

1550, TITLE: All-in-One Image Restoration for Unknown Corruption
http://openaccess.thecvf.com//content/CVPR2022/html/Li_All-in-One_Image_Restoration_for_Unknown_Corruption_CVPR_2022_paper.html
AUTHORS: Boyun Li, Xiao Liu, Peng Hu, Zhongqin Wu, Jiancheng Lv, Xi Peng
HIGHLIGHT: In this paper, we study a challenging problem in image restoration, namely, how to develop an all-in-one method that could recover images from a variety of unknown corruption types and levels.

1551, TITLE: Syntax-Aware Network for Handwritten Mathematical Expression Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Yuan_Syntax-Aware_Network_for_Handwritten_Mathematical_Expression_Recognition_CVPR_2022_paper.html
AUTHORS: Ye Yuan, Xiao Liu, Wondimu Dikubab, Hui Liu, Zhilong Ji, Zhongqin Wu, Xiang Bai
HIGHLIGHT: In this paper, we propose a simple and efficient method for HMER, which is the first to incorporate syntax information into an encoder-decoder network.

1552, TITLE: Sketching Without Worrying: Noise-Tolerant Sketch-Based Image Retrieval
http://openaccess.thecvf.com//content/CVPR2022/html/Bhunia_Sketching_Without_Worrying_Noise-Tolerant_Sketch-Based_Image_Retrieval_CVPR_2022_paper.html
AUTHORS: Ayan Kumar Bhunia, Subhadeep Koley, Abdullah Faiz Ur Rahman Khilji, Aneeshan Sain, Pinaki Nath Chowdhury, Tao Xiang, Yi-Zhe Song
HIGHLIGHT: Sketching enables many exciting applications, notably, image retrieval. The fear-to-sketch problem (i.e., "I can't sketch") has however proven to be fatal for its widespread adoption. This paper tackles this "fear" head on, and for the first time, proposes an auxiliary module for existing retrieval models that predominantly lets the users sketch without having to worry.

1553, TITLE: PUMP: Pyramidal and Uniqueness Matching Priors for Unsupervised Learning of Local Descriptors
http://openaccess.thecvf.com//content/CVPR2022/html/Revaud_PUMP_Pyramidal_and_Uniqueness_Matching_Priors_for_Unsupervised_Learning_of_CVPR_2022_paper.html
AUTHORS: Jérome Revaud, Vincent Leroy, Philippe Weinzaepfel, Boris Chidlovskii
HIGHLIGHT: In this paper, we propose to explicitly integrate two matching priors in a single loss in order to learn local descriptors without supervision.

1554, TITLE: PlanarRecon: Real-Time 3D Plane Detection and Reconstruction From Posed Monocular Videos
http://openaccess.thecvf.com//content/CVPR2022/html/Xie_PlanarRecon_Real-Time_3D_Plane_Detection_and_Reconstruction_From_Posed_Monocular_CVPR_2022_paper.html
AUTHORS: Yiming Xie, Matheus Gadelha, Fengting Yang, Xiaowei Zhou, Huaizu Jiang
HIGHLIGHT: We present PlanarRecon -- a novel framework for globally coherent detection and reconstruction of 3D planes from a posed monocular video.

1555, TITLE: Deep Equilibrium Optical Flow Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Bai_Deep_Equilibrium_Optical_Flow_Estimation_CVPR_2022_paper.html
AUTHORS: Shaojie Bai, Zhengyang Geng, Yash Savani, J. Zico Kolter
HIGHLIGHT: They can converge poorly and thereby suffer from performance degradation. To combat these drawbacks, we propose deep equilibrium (DEQ) flow estimators, an approach that directly solves for the flow as the infinite-level fixed point of an implicit layer (using any black-box solver), and differentiates through this fixed point analytically (thus requiring O(1) training memory).

1556, TITLE: Optimizing Video Prediction via Video Frame Interpolation
http://openaccess.thecvf.com//content/CVPR2022/html/Wu_Optimizing_Video_Prediction_via_Video_Frame_Interpolation_CVPR_2022_paper.html
AUTHORS: Yue Wu, Qiang Wen, Qifeng Chen
HIGHLIGHT: Inspired by the photo-realistic results of video frame interpolation, we present a new optimization framework for video prediction via video frame interpolation, in which we solve an extrapolation problem based on an interpolation model.

1557, TITLE: Motron: Multimodal Probabilistic Human Motion Forecasting
http://openaccess.thecvf.com//content/CVPR2022/html/Salzmann_Motron_Multimodal_Probabilistic_Human_Motion_Forecasting_CVPR_2022_paper.html
AUTHORS: Tim Salzmann, Marco Pavone, Markus Ryll

HIGHLIGHT: We present Motron, a multimodal, probabilistic, graph-structured model, that captures human's multimodality using probabilistic methods while being able to output deterministic maximum-likelihood motions and corresponding confidence values for each mode.

1558, TITLE: Episodic Memory Question Answering
http://openaccess.thecvf.com//content/CVPR2022/html/Datta_Episodic_Memory_Question_Answering_CVPR_2022_paper.html
AUTHORS: Samyak Datta, Sameer Dharur, Vincent Cartillier, Ruta Desai, Mukul Khanna, Dhruv Batra, Devi Parikh
HIGHLIGHT: Towards that end, we introduce (1) a new task -- Episodic Memory Question Answering (EMQA) wherein an egocentric AI assistant is provided with a video sequence (the tour) and a question as an input and is asked to localize its answer to the question within the tour, (2) a dataset of grounded questions designed to probe the agent's spatio-temporal understanding of the tour, and (3) a model for the task that encodes the scene as an allocentric, top-down semantic feature map and grounds the question into the map to localize the answer.

1559, TITLE: Continual Stereo Matching of Continuous Driving Scenes With Growing Architecture
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Continual_Stereo_Matching_of_Continuous_Driving_Scenes_With_Growing_Architecture_CVPR_2022_paper.html
AUTHORS: Chenghao Zhang, Kun Tian, Bin Fan, Gaofeng Meng, Zhaoxiang Zhang, Chunhong Pan
HIGHLIGHT: To address these challenges, we propose to perform continual stereo matching where a model is tasked to 1) continually learn new scenes, 2) overcome forgetting previously learned scenes, and 3) continuously predict disparities at deployment. We achieve this goal by introducing a Reusable Architecture Growth (RAG) framework.

1560, TITLE: Few-Shot Backdoor Defense Using Shapley Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Guan_Few-Shot_Backdoor_Defense_Using_Shapley_Estimation_CVPR_2022_paper.html
AUTHORS: Jiyang Guan, Zhuozhuo Tu, Ran He, Dacheng Tao
HIGHLIGHT: To determine the triggered neurons and protect against backdoor attacks, we exploit Shapley value and develop a new approach called Shapley Pruning (ShapPruning) that successfully mitigates backdoor attacks from models in a data-insufficient situation (1 image per class or even free of data).

1561, TITLE: Cycle-Consistent Counterfactuals by Latent Transformations
http://openaccess.thecvf.com//content/CVPR2022/html/Khorram_Cycle-Consistent_Counterfactuals_by_Latent_Transformations_CVPR_2022_paper.html
AUTHORS: Saeed Khorram, Li Fuxin
HIGHLIGHT: We propose a novel approach, Cycle-Consistent Counterfactuals by Latent Transformations (C3LT), which learns a latent transformation that automatically generates visual CFs by steering in the latent space of generative models.

1562, TITLE: ADeLA: Automatic Dense Labeling With Attention for Viewpoint Shift in Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Ren_ADeLA_Automatic_Dense_Labeling_With_Attention_for_Viewpoint_Shift_in_CVPR_2022_paper.html
AUTHORS: Hanxiang Ren, Yanchao Yang, He Wang, Bokui Shen, Qingnan Fan, Youyi Zheng, C. Karen Liu, Leonidas J. Guibas
HIGHLIGHT: We describe a method to deal with performance drop in semantic segmentation caused by viewpoint changes within multi-camera systems, where temporally paired images are readily available, but the annotations may only be abundant for a few typical views.

1563, TITLE: Joint Hand Motion and Interaction Hotspots Prediction From Egocentric Videos
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Joint_Hand_Motion_and_Interaction_Hotspots_Prediction_From_Egocentric_Videos_CVPR_2022_paper.html
AUTHORS: Shaowei Liu, Subarna Tripathi, Somdeb Majumdar, Xiaolong Wang
HIGHLIGHT: We propose to forecast future hand-object interactions given an egocentric video.

1564, TITLE: Blind Face Restoration via Integrating Face Shape and Generative Priors
http://openaccess.thecvf.com//content/CVPR2022/html/Zhu_Blind_Face_Restoration_via_Integrating_Face_Shape_and_Generative_Priors_CVPR_2022_paper.html
AUTHORS: Feida Zhu, Junwei Zhu, Wenqing Chu, Xinyi Zhang, Xiaozhong Ji, Chengjie Wang, Ying Tai
HIGHLIGHT: In this work, we propose to integrate shape and generative priors to guide the challenging blind face restoration.

1565, TITLE: MixSTE: Seq2seq Mixed Spatio-Temporal Encoder for 3D Human Pose Estimation in Video
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_MixSTE_Seq2seq_Mixed_Spatio-Temporal_Encoder_for_3D_Human_Pose_Estimation_CVPR_2022_paper.html
AUTHORS: Jinlu Zhang, Zhigang Tu, Jianyu Yang, Yujin Chen, Junsong Yuan
HIGHLIGHT: We propose MixSTE (Mixed Spatio-Temporal Encoder), which has a temporal transformer block to separately model the temporal motion of each joint and a spatial transformer block to learn inter-joint spatial correlation.

1566, TITLE: Safe-Student for Safe Deep Semi-Supervised Learning With Unseen-Class Unlabeled Data
http://openaccess.thecvf.com//content/CVPR2022/html/He_Safe-Student_for_Safe_Deep_Semi-Supervised_Learning_With_Unseen-Class_Unlabeled_Data_CVPR_2022_paper.html
AUTHORS: Rundong He, Zhongyi Han, Xiankai Lu, Yilong Yin
HIGHLIGHT: In this paper, we consider the problem of safe SSL scenario where unseen-class instances appear in the unlabeled data.

1567, TITLE:	Learning To Zoom Inside Camera Imaging Pipeline
http://openaccess.thecvf.com//content/CVPR2022/html/Tang_Learning_To_Zoom_Inside_Camera_Imaging_Pipeline_CVPR_2022_paper.html
AUTHORS:	Chengzhou Tang, Yuqiang Yang, Bing Zeng, Ping Tan, Shuaicheng Liu
HIGHLIGHT:	This paper proposes to zoom an image from RAW to RAW inside the camera imaging pipeline.

1568, TITLE:	High-Fidelity GAN Inversion for Image Attribute Editing
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_High-Fidelity_GAN_Inversion_for_Image_Attribute_Editing_CVPR_2022_paper.html
AUTHORS:	Tengfei Wang, Yong Zhang, Yanbo Fan, Jue Wang, Qifeng Chen
HIGHLIGHT:	We present a novel high-fidelity generative adversarial network (GAN) inversion framework that enables attribute editing with image-specific details well-preserved (e.g., background, appearance, and illumination).

1569, TITLE:	RCP: Recurrent Closest Point for Point Cloud
http://openaccess.thecvf.com//content/CVPR2022/html/Gu_RCP_Recurrent_Closest_Point_for_Point_Cloud_CVPR_2022_paper.html
AUTHORS:	Xiaodong Gu, Chengzhou Tang, Weihao Yuan, Zuozhuo Dai, Siyu Zhu, Ping Tan
HIGHLIGHT:	However, these methods are limited by the fact that it is difficult to define a search window on point clouds because of the irregular data structure. In this paper, we avoid this irregularity by a simple yet effective method.

1570, TITLE:	gDNA: Towards Generative Detailed Neural Avatars
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_gDNA_Towards_Generative_Detailed_Neural_Avatars_CVPR_2022_paper.html
AUTHORS:	Xu Chen, Tianjian Jiang, Jie Song, Jinlong Yang, Michael J. Black, Andreas Geiger, Otmar Hilliges
HIGHLIGHT:	In this paper, we propose a novel method that learns to generate detailed 3D shapes of people in a variety of garments with corresponding skinning weights.

1571, TITLE:	A Dual Weighting Label Assignment Scheme for Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Li_A_Dual_Weighting_Label_Assignment_Scheme_for_Object_Detection_CVPR_2022_paper.html
AUTHORS:	Shuai Li, Chenhang He, Ruihuang Li, Lei Zhang
HIGHLIGHT:	In this paper, we explore a new weighting paradigm, termed dual weighting (DW), to specify pos and neg weights separately.

1572, TITLE:	FAM: Visual Explanations for the Feature Representations From Deep Convolutional Networks
http://openaccess.thecvf.com//content/CVPR2022/html/Wu_FAM_Visual_Explanations_for_the_Feature_Representations_From_Deep_Convolutional_CVPR_2022_paper.html
AUTHORS:	Yuxi Wu, Changhuai Chen, Jun Che, Shiliang Pu
HIGHLIGHT:	In this case, employing an existing class or the similarity with other image is unable to provide a complete and reliable visual explanation. To handle this task, we propose a novel visual explanation paradigm called Feature Activation Mapping (FAM) in this paper.

1573, TITLE:	Hyperbolic Vision Transformers: Combining Improvements in Metric Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Ermolov_Hyperbolic_Vision_Transformers_Combining_Improvements_in_Metric_Learning_CVPR_2022_paper.html
AUTHORS:	Aleksandr Ermolov, Leyla Mirvakhabova, Valentin Khrulkov, Nicu Sebe, Ivan Oseledets
HIGHLIGHT:	An emerging interest in learning hyperbolic data embeddings suggests that hyperbolic geometry can be beneficial for natural data. Following this line of work, we propose a new hyperbolic-based model for metric learning.

1574, TITLE:	MaskGIT: Masked Generative Image Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Chang_MaskGIT_Masked_Generative_Image_Transformer_CVPR_2022_paper.html
AUTHORS:	Huiwen Chang, Han Zhang, Lu Jiang, Ce Liu, William T. Freeman
HIGHLIGHT:	This paper proposes a novel image synthesis paradigm using a bidirectional transformer decoder, which we term MaskGIT.

1575, TITLE:	Revisiting the "Video" in Video-Language Understanding
http://openaccess.thecvf.com//content/CVPR2022/html/Buch_Revisiting_the_Video_in_Video-Language_Understanding_CVPR_2022_paper.html
AUTHORS:	Shyamal Buch, Cristóbal Eyzaguirre, Adrien Gaidon, Jiajun Wu, Li Fei-Fei, Juan Carlos Niebles
HIGHLIGHT:	We propose the atemporal probe (ATP), a new model for video-language analysis which provides a stronger bound on the baseline accuracy of multimodal models constrained by image-level understanding.

1576, TITLE:	Local Texture Estimator for Implicit Representation Function
http://openaccess.thecvf.com//content/CVPR2022/html/Lee_Local_Texture_Estimator_for_Implicit_Representation_Function_CVPR_2022_paper.html
AUTHORS:	Jaewon Lee, Kyong Hwan Jin
HIGHLIGHT:	In this paper, we propose a Local Texture Estimator (LTE), a dominant-frequency estimator for natural images, enabling an implicit function to capture fine details while reconstructing images in a continuous manner.

1577, TITLE: Instance-Aware Dynamic Neural Network Quantization
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Instance-
Aware_Dynamic_Neural_Network_Quantization_CVPR_2022_paper.html
AUTHORS: Zhenhua Liu, Yunhe Wang, Kai Han, Siwei Ma, Wen Gao
HIGHLIGHT: In this paper, we present to conduct the low-bit quantization for each image individually, and develop a dynamic quantization scheme for exploring their optimal bit-widths.

1578, TITLE: When To Prune? A Policy Towards Early Structural Pruning
http://openaccess.thecvf.com//content/CVPR2022/html/Shen_When_To_Prune_A_Policy_Towards_Early_Structural_Pruning_CVPR_2022_paper.html
AUTHORS: Maying Shen, Pavlo Molchanov, Hongxu Yin, Jose M. Alvarez
HIGHLIGHT: We attempt to combine the benefits of both directions and propose a policy that prunes as early as possible during training without hurting performance.

1579, TITLE: COTS: Collaborative Two-Stream Vision-Language Pre-Training Model for Cross-Modal Retrieval
http://openaccess.thecvf.com//content/CVPR2022/html/Lu_COTS_Collaborative_Two-Stream_Vision-Language_Pre-Training_Model_for_Cross-Modal_Retrieval_CVPR_2022_paper.html
AUTHORS: Haoyu Lu, Nanyi Fei, Yuqi Huo, Yizhao Gao, Zhiwu Lu, Ji-Rong Wen
HIGHLIGHT: Recently, two-stream methods like CLIP and ALIGN with high inference efficiency have also shown promising performance, however, they only consider instance-level alignment between the two streams (thus there is still room for improvement). To overcome these limitations, we propose a novel COllaborative Two-Stream vision-language pre-training model termed COTS for image-text retrieval by enhancing cross-modal interaction.

1580, TITLE: Degree-of-Linear-Polarization-Based Color Constancy
http://openaccess.thecvf.com//content/CVPR2022/html/Ono_Degree-of-Linear-Polarization-Based_Color_Constancy_CVPR_2022_paper.html
AUTHORS: Taishi Ono, Yuhi Kondo, Legong Sun, Teppei Kurita, Yusuke Moriuchi
HIGHLIGHT: This paper shows that the degree of linear polarization dramatically solves the color constancy problem because it allows us to find achromatic pixels stably.

1581, TITLE: A Voxel Graph CNN for Object Classification With Event Cameras
http://openaccess.thecvf.com//content/CVPR2022/html/Deng_A_Voxel_Graph_CNN_for_Object_Classification_With_Event_Cameras_CVPR_2022_paper.html
AUTHORS: Yongjian Deng, Hao Chen, Hai Liu, Youfu Li
HIGHLIGHT: This study aims to address the core problem of balancing accuracy and model complexity for event-based classification models.

1582, TITLE: On the Importance of Asymmetry for Siamese Representation Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_On_the_Importance_of_Asymmetry_for_Siamese_Representation_Learning_CVPR_2022_paper.html
AUTHORS: Xiao Wang, Haoqi Fan, Yuandong Tian, Daisuke Kihara, Xinlei Chen
HIGHLIGHT: In this work, we conduct a formal study on the importance of asymmetry by explicitly distinguishing the two encoders within the network -- one produces source encodings and the other targets.

1583, TITLE: Probing Representation Forgetting in Supervised and Unsupervised Continual Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Davari_Probing_Representation_Forgetting_in_Supervised_and_Unsupervised_Continual_Learning_CVPR_2022_paper.html
AUTHORS: MohammadReza Davari, Nader Asadi, Sudhir Mudur, Rahaf Aljundi, Eugene Belilovsky
HIGHLIGHT: In this work we consider the concept of representation forgetting, observed by using the difference in performance of an optimal linear classifier before and after a new task is introduced.

1584, TITLE: ViSTA: Vision and Scene Text Aggregation for Cross-Modal Retrieval
http://openaccess.thecvf.com//content/CVPR2022/html/Cheng_ViSTA_Vision_and_Scene_Text_Aggregation_for_Cross-Modal_Retrieval_CVPR_2022_paper.html
AUTHORS: Mengjun Cheng, Yipeng Sun, Longchao Wang, Xiongwei Zhu, Kun Yao, Jie Chen, Guoli Song, Junyu Han, Jingtuo Liu, Errui Ding, Jingdong Wang
HIGHLIGHT: To tackle the modality missing problem of scene text, we propose a novel fusion token based transformer aggregation approach to exchange the necessary scene text information only through the fusion token and concentrate on the most important features in each modality.

1585, TITLE: DenseCLIP: Language-Guided Dense Prediction With Context-Aware Prompting
http://openaccess.thecvf.com//content/CVPR2022/html/Rao_DenseCLIP_Language-Guided_Dense_Prediction_With_Context-Aware_Prompting_CVPR_2022_paper.html
AUTHORS: Yongming Rao, Wenliang Zhao, Guangyi Chen, Yansong Tang, Zheng Zhu, Guan Huang, Jie Zhou, Jiwen Lu
HIGHLIGHT: In this work, we present a new framework for dense prediction by implicitly and explicitly leveraging the pre-trained knowledge from CLIP.

1586, TITLE: Exploring Effective Data for Surrogate Training Towards Black-Box Attack

http://openaccess.thecvf.com//content/CVPR2022/html/Sun_Exploring_Effective_Data_for_Surrogate_Training_Towards_Black-Box_Attack_CVPR_2022_paper.html
AUTHORS: Xuxiang Sun, Gong Cheng, Hongda Li, Lei Pei, Junwei Han
HIGHLIGHT: To this end, we propose a triple-player framework by introducing a discriminator into the traditional data-free framework.

1587, TITLE: JRDB-Act: A Large-Scale Dataset for Spatio-Temporal Action, Social Group and Activity Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Ehsanpour_JRDB-Act_A_Large-Scale_Dataset_for_Spatio-Temporal_Action_Social_Group_and_CVPR_2022_paper.html
AUTHORS: Mahsa Ehsanpour, Fatemeh Saleh, Silvio Savarese, Ian Reid, Hamid Rezatofighi
HIGHLIGHT: In this paper, we introduce JRDB-Act, as an extension of the existing JRDB, which is captured by a social mobile manipulator and reflects a real distribution of human daily-life actions in a university campus environment.

1588, TITLE: AR-NeRF: Unsupervised Learning of Depth and Defocus Effects From Natural Images With Aperture Rendering Neural Radiance Fields
http://openaccess.thecvf.com//content/CVPR2022/html/Kaneko_AR-NeRF_Unsupervised_Learning_of_Depth_and_Defocus_Effects_From_Natural_CVPR_2022_paper.html
AUTHORS: Takuhiro Kaneko
HIGHLIGHT: As an alternative to an AR-GAN, we propose an aperture rendering NeRF (AR-NeRF), which can utilize viewpoint and defocus cues in a unified manner by representing both factors in a common ray-tracing framework.

1589, TITLE: Likert Scoring With Grade Decoupling for Long-Term Action Assessment
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_Likert_Scoring_With_Grade_Decoupling_for_Long-Term_Action_Assessment_CVPR_2022_paper.html
AUTHORS: Angchi Xu, Ling-An Zeng, Wei-Shi Zheng
HIGHLIGHT: Hence, the final score should be determined by the comprehensive effect of different grades exhibited in the video. To explore this latent relationship, we design a novel Likert scoring paradigm inspired by the Likert scale in psychometrics, in which we quantify the grades explicitly and generate the final quality score by combining the quantitative values and the corresponding responses estimated from the video, instead of performing direct regression.

1590, TITLE: Many-to-Many Splatting for Efficient Video Frame Interpolation
http://openaccess.thecvf.com//content/CVPR2022/html/Hu_Many-to-Many_Splatting_for_Efficient_Video_Frame_Interpolation_CVPR_2022_paper.html
AUTHORS: Ping Hu, Simon Niklaus, Stan Sclaroff, Kate Saenko
HIGHLIGHT: In this work, we propose a fully differentiable Many-to-Many (M2M) splatting framework to interpolate frames efficiently.

1591, TITLE: Investigating Top-k White-Box and Transferable Black-Box Attack
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Investigating_Top-k_White-Box_and_Transferable_Black-Box_Attack_CVPR_2022_paper.html
AUTHORS: Chaoning Zhang, Philipp Benz, Adil Karjauv, Jae Won Cho, Kang Zhang, In So Kweon
HIGHLIGHT: To this end, we propose a new normalized CE loss that guides the logit to be updated in the direction of implicitly maximizing its rank distance from the ground-truth class.

1592, TITLE: Decoupling and Recoupling Spatiotemporal Representation for RGB-D-Based Motion Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Zhou_Decoupling_and_Recoupling_Spatiotemporal_Representation_for_RGB-D-Based_Motion_Recognition_CVPR_2022_paper.html
AUTHORS: Benjia Zhou, Pichao Wang, Jun Wan, Yanyan Liang, Fan Wang, Du Zhang, Zhen Lei, Hao Li, Rong Jin
HIGHLIGHT: Although previous RGB-D-based motion recognition methods have achieved promising performance through the tightly coupled multi-modal spatiotemporal representation, they still suffer from (i) optimization difficulty under small data setting due to the tightly spatiotemporal-entangled modeling; (ii) information redundancy as it usually contains lots of marginal information that is weakly relevant to classification; and (iii) low interaction between multi-modal spatiotemporal information caused by insufficient late fusion. To alleviate these drawbacks, we propose to decouple and recouple spatiotemporal representation for RGB-D-based motion recognition.

1593, TITLE: Learning To Learn by Jointly Optimizing Neural Architecture and Weights
http://openaccess.thecvf.com//content/CVPR2022/html/Ding_Learning_To_Learn_by_Jointly_Optimizing_Neural_Architecture_and_Weights_CVPR_2022_paper.html
AUTHORS: Yadong Ding, Yu Wu, Chengyue Huang, Siliang Tang, Yi Yang, Longhui Wei, Yueting Zhuang, Qi Tian
HIGHLIGHT: In this paper, we aim to obtain better meta-learners by co-optimizing the architecture and meta-weights simultaneously.

1594, TITLE: Attributable Visual Similarity Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Attributable_Visual_Similarity_Learning_CVPR_2022_paper.html
AUTHORS: Borui Zhang, Wenzhao Zheng, Jie Zhou, Jiwen Lu
HIGHLIGHT: This paper proposes an attributable visual similarity learning (AVSL) framework for a more accurate and explainable similarity measure between images.

1595, TITLE: A Self-Supervised Descriptor for Image Copy Detection

http://openaccess.thecvf.com//content/CVPR2022/html/Pizzi_A_Self-Supervised_Descriptor_for_Image_Copy_Detection_CVPR_2022_paper.html
AUTHORS: Ed Pizzi, Sreya Dutta Roy, Sugosh Nagavara Ravindra, Priya Goyal, Matthijs Douze
HIGHLIGHT: We introduce SSCD, a model that builds on a recent self-supervised contrastive training objective.

1596, TITLE: DyTox: Transformers for Continual Learning With DYnamic TOken eXpansion
http://openaccess.thecvf.com//content/CVPR2022/html/Douillard_DyTox_Transformers_for_Continual_Learning_With_DYnamic_TOken_eXpansion_CVPR_2022_paper.html
AUTHORS: Arthur Douillard, Alexandre Ramé, Guillaume Couairon, Matthieu Cord
HIGHLIGHT: In this paper, we propose a transformer architecture based on a dedicated encoder/decoder framework.

1597, TITLE: Towards Robust and Adaptive Motion Forecasting: A Causal Representation Perspective
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Towards_Robust_and_Adaptive_Motion_Forecasting_A_Causal_Representation_Perspective_CVPR_2022_paper.html
AUTHORS: Yuejiang Liu, Riccardo Cadei, Jonas Schweizer, Sherwin Bahmani, Alexandre Alahi
HIGHLIGHT: Yet, the current paradigm suffers from two shortcomings: brittle under distribution shifts and inefficient for knowledge transfer. In this work, we propose to address these challenges from a causal representation perspective.

1598, TITLE: Manifold Learning Benefits GANs
http://openaccess.thecvf.com//content/CVPR2022/html/Ni_Manifold_Learning_Benefits_GANs_CVPR_2022_paper.html
AUTHORS: Yao Ni, Piotr Koniusz, Richard Hartley, Richard Nock
HIGHLIGHT: In this paper, we improve Generative Adversarial Networks by incorporating a manifold learning step into the discriminator.

1599, TITLE: A Keypoint-Based Global Association Network for Lane Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_A_Keypoint-Based_Global_Association_Network_for_Lane_Detection_CVPR_2022_paper.html
AUTHORS: Jinsheng Wang, Yinchao Ma, Shaofei Huang, Tianrui Hui, Fei Wang, Chen Qian, Tianzhu Zhang
HIGHLIGHT: In this paper, we propose a Global Association Network (GANet) to formulate the lane detection problem from a new perspective, where each keypoint is directly regressed to the starting point of the lane line instead of point-by-point extension.

1600, TITLE: Negative-Aware Attention Framework for Image-Text Matching
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Negative-Aware_Attention_Framework_for_Image-Text_Matching_CVPR_2022_paper.html
AUTHORS: Kun Zhang, Zhendong Mao, Quan Wang, Yongdong Zhang
HIGHLIGHT: We thereby propose a novel Negative-Aware Attention Framework (NAAF), which explicitly exploits both the positive effect of matched fragments and the negative effect of mismatched fragments to jointly infer image-text similarity.

1601, TITLE: Semantic-Aligned Fusion Transformer for One-Shot Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Zhao_Semantic-Aligned_Fusion_Transformer_for_One-Shot_Object_Detection_CVPR_2022_paper.html
AUTHORS: Yizhou Zhao, Xun Guo, Yan Lu
HIGHLIGHT: Yet, their performances are often unsatisfactory. In this paper, we attribute this to inappropriate correlation methods that misalign query-support semantics by overlooking spatial structures and scale variances.

1602, TITLE: Beyond Supervised vs. Unsupervised: Representative Benchmarking and Analysis of Image Representation Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Gwilliam_Beyond_Supervised_vs._Unsupervised_Representative_Benchmarking_and_Analysis_of_Image_CVPR_2022_paper.html
AUTHORS: Matthew Gwilliam, Abhinav Shrivastava
HIGHLIGHT: In this paper, we compare methods using performance-based benchmarks such as linear evaluation, nearest neighbor classification, and clustering for several different datasets, demonstrating the lack of a clear frontrunner within the current state-of-the-art.

1603, TITLE: Few-Shot Incremental Learning for Label-to-Image Translation
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Few-Shot_Incremental_Learning_for_Label-to-Image_Translation_CVPR_2022_paper.html
AUTHORS: Pei Chen, Yangkang Zhang, Zejian Li, Lingyun Sun
HIGHLIGHT: In this paper, we introduce a few-shot incremental learning method for label-to-image translation.

1604, TITLE: Discrete Time Convolution for Fast Event-Based Stereo
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Discrete_Time_Convolution_for_Fast_Event-Based_Stereo_CVPR_2022_paper.html
AUTHORS: Kaixuan Zhang, Kaiwei Che, Jianguo Zhang, Jie Cheng, Ziyang Zhang, Qinghai Guo, Luziwei Leng
HIGHLIGHT: Inspired by continuous dynamics of biological neuron models, we propose a novel encoding method for sparse events - continuous time convolution (CTC) - which learns to model the spatial feature of the data with intrinsic dynamics.

1605, TITLE: An Image Patch Is a Wave: Phase-Aware Vision MLP
http://openaccess.thecvf.com//content/CVPR2022/html/Tang_An_Image_Patch_Is_a_Wave_Phase-Aware_Vision_MLP_CVPR_2022_paper.html

AUTHORS: Yehui Tang, Kai Han, Jianyuan Guo, Chang Xu, Yanxi Li, Chao Xu, Yunhe Wang
HIGHLIGHT: To dynamically aggregate tokens, we propose to represent each token as a wave function with two parts, amplitude and phase.

1606, TITLE: Escaping Data Scarcity for High-Resolution Heterogeneous Face Hallucination
http://openaccess.thecvf.com//content/CVPR2022/html/Mei_Escaping_Data_Scarcity_for_High-Resolution_Heterogeneous_Face_Hallucination_CVPR_2022_paper.html
AUTHORS: Yiqun Mei, Pengfei Guo, Vishal M. Patel
HIGHLIGHT: In this paper, we propose a new face hallucination paradigm for HFR, which not only enables data-efficient synthesis but also allows to scale up model training without breaking any privacy policy.

1607, TITLE: Visual Acoustic Matching
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Visual_Acoustic_Matching_CVPR_2022_paper.html
AUTHORS: Changan Chen, Ruohan Gao, Paul Calamia, Kristen Grauman
HIGHLIGHT: Given an image of the target environment and a waveform for the source audio, the goal is to re-synthesize the audio to match the target room acoustics as suggested by its visible geometry and materials. To address this novel task, we propose a cross-modal transformer model that uses audio-visual attention to inject visual properties into the audio and generate realistic audio output.

1608, TITLE: Shunted Self-Attention via Multi-Scale Token Aggregation
http://openaccess.thecvf.com//content/CVPR2022/html/Ren_Shunted_Self-Attention_via_Multi-Scale_Token_Aggregation_CVPR_2022_paper.html
AUTHORS: Sucheng Ren, Daquan Zhou, Shengfeng He, Jiashi Feng, Xinchao Wang
HIGHLIGHT: Such a constraint inevitably limits the ability of each self-attention layer in capturing multi-scale features, thereby leading to performance degradation in handling images with multiple objects of different scales. To address this issue, we propose a novel and generic strategy, termed shunted self-attention (SSA), that allows ViTs to model the attentions at hybrid scales per attention layer.

1609, TITLE: Shadows Can Be Dangerous: Stealthy and Effective Physical-World Adversarial Attack by Natural Phenomenon
http://openaccess.thecvf.com//content/CVPR2022/html/Zhong_Shadows_Can_Be_Dangerous_Stealthy_and_Effective_Physical-World_Adversarial_Attack_CVPR_2022_paper.html
AUTHORS: Yiqi Zhong, Xianming Liu, Deming Zhai, Junjun Jiang, Xiangyang Ji
HIGHLIGHT: In this paper, we study a new type of optical adversarial examples, in which the perturbations are generated by a very common natural phenomenon, shadow, to achieve naturalistic and stealthy physical-world adversarial attack under the black-box setting.

1610, TITLE: ImplicitAtlas: Learning Deformable Shape Templates in Medical Imaging
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_ImplicitAtlas_Learning_Deformable_Shape_Templates_in_Medical_Imaging_CVPR_2022_paper.html
AUTHORS: Jiancheng Yang, Udaranga Wickramasinghe, Bingbing Ni, Pascal Fua
HIGHLIGHT: This is in part because large training databases do not exist and in part because biomedical annotations are often noisy. In this paper, we show that by introducing templates within the deep learning pipeline we can overcome these problems.

1611, TITLE: Unified Multivariate Gaussian Mixture for Efficient Neural Image Compression
http://openaccess.thecvf.com//content/CVPR2022/html/Zhu_Unified_Multivariate_Gaussian_Mixture_for_Efficient_Neural_Image_Compression_CVPR_2022_paper.html
AUTHORS: Xiaosu Zhu, Jingkuan Song, Lianli Gao, Feng Zheng, Heng Tao Shen
HIGHLIGHT: Specifically, a multivariate Gaussian mixture is proposed with means and covariances to be estimated. Then, a novel probabilistic vector quantization is utilized to effectively approximate means, and remaining covariances are further induced to a unified mixture and solved by cascaded estimation without context models involved.

1612, TITLE: 3D Photo Stylization: Learning To Generate Stylized Novel Views From a Single Image
http://openaccess.thecvf.com//content/CVPR2022/html/Mu_3D_Photo_Stylization_Learning_To_Generate_Stylized_Novel_Views_From_CVPR_2022_paper.html
AUTHORS: Fangzhou Mu, Jian Wang, Yicheng Wu, Yin Li
HIGHLIGHT: In this paper, we make a connection between the two, and address the challenging task of 3D photo stylization - generating stylized novel views from a single image given an arbitrary style.

1613, TITLE: Improving Visual Grounding With Visual-Linguistic Verification and Iterative Reasoning
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Improving_Visual_Grounding_With_Visual-Linguistic_Verification_and_Iterative_Reasoning_CVPR_2022_paper.html
AUTHORS: Li Yang, Yan Xu, Chunfeng Yuan, Wei Liu, Bing Li, Weiming Hu
HIGHLIGHT: In this paper, we propose a transformer-based framework for accurate visual grounding by establishing text-conditioned discriminative features and performing multi-stage cross-modal reasoning.

1614, TITLE: Contrastive Learning for Space-Time Correspondence via Self-Cycle Consistency
http://openaccess.thecvf.com//content/CVPR2022/html/Son_Contrastive_Learning_for_Space-Time_Correspondence_via_Self-Cycle_Consistency_CVPR_2022_paper.html
AUTHORS: Jeany Son

HIGHLIGHT: We propose a novel probabilistic method employing Bayesian Model Averaging and self-cycle regularization for spatio-temporal correspondence learning in videos within a self-supervised learning framework.

1615, TITLE: Learning Robust Image-Based Rendering on Sparse Scene Geometry via Depth Completion
http://openaccess.thecvf.com//content/CVPR2022/html/Sun_Learning_Robust_Image-Based_Rendering_on_Sparse_Scene_Geometry_via_Depth_CVPR_2022_paper.html
AUTHORS: Yuqi Sun, Shili Zhou, Ri Cheng, Weimin Tan, Bo Yan, Lang Fu
HIGHLIGHT: When only few views are provided, the performance of these methods drops off significantly, as the scene geometry becomes sparse as well. Therefore, in this paper, we propose Sparse-IBRNet (SIBRNet) to perform robust IBR on sparse scene geometry by depth completion.

1616, TITLE: Scale-Equivalent Distillation for Semi-Supervised Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Guo_Scale-Equivalent_Distillation_for_Semi-Supervised_Object_Detection_CVPR_2022_paper.html
AUTHORS: Qiushan Guo, Yao Mu, Jianyu Chen, Tianqi Wang, Yizhou Yu, Ping Luo
HIGHLIGHT: Further, we overcome these challenges by introducing a novel approach, Scale-Equivalent Distillation (SED), which is a simple yet effective end-to-end knowledge distillation framework robust to large object size variance and class imbalance.

1617, TITLE: Recurrent Variational Network: A Deep Learning Inverse Problem Solver Applied to the Task of Accelerated MRI Reconstruction
http://openaccess.thecvf.com//content/CVPR2022/html/Yiasemis_Recurrent_Variational_Network_A_Deep_Learning_Inverse_Problem_Solver_Applied_CVPR_2022_paper.html
AUTHORS: George Yiasemis, Jan-Jakob Sonke, Clarisa Sánchez, Jonas Teuwen
HIGHLIGHT: In this work we present a novel Deep Learning-based Inverse Problem solver applied to the task of Accelerated MRI Reconstruction, called the Recurrent Variational Network (RecurrentVarNet), by exploiting the properties of Convolutional Recurrent Neural Networks and unrolled algorithms for solving Inverse Problems.

1618, TITLE: SelfD: Self-Learning Large-Scale Driving Policies From the Web
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_SelfD_Self-Learning_Large-Scale_Driving_Policies_From_the_Web_CVPR_2022_paper.html
AUTHORS: Jimuyang Zhang, Ruizhao Zhu, Eshed Ohn-Bar
HIGHLIGHT: In this work, we introduce SelfD, a framework for learning scalable driving by utilizing large amounts of online monocular images.

1619, TITLE: "The Pedestrian Next to the Lamppost" Adaptive Object Graphs for Better Instantaneous Mapping
http://openaccess.thecvf.com//content/CVPR2022/html/Saha_The_Pedestrian_Next_to_the_Lamppost_Adaptive_Object_Graphs_for_CVPR_2022_paper.html
AUTHORS: Avishkar Saha, Oscar Mendez, Chris Russell, Richard Bowden
HIGHLIGHT: While such an increase in error is entirely expected - localization is harder at distance - much of the drop in performance can be attributed to the cues used by current texture-based models, in particular, they make heavy use of object-ground intersections (such as shadows), which become increasingly sparse and uncertain for distant objects. In this work, we address these shortcomings in BEV-mapping by learning the spatial relationship between objects in a scene.

1620, TITLE: Attribute Group Editing for Reliable Few-Shot Image Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Ding_Attribute_Group_Editing_for_Reliable_Few-Shot_Image_Generation_CVPR_2022_paper.html
AUTHORS: Guanqi Ding, Xinzhe Han, Shuhui Wang, Shuzhe Wu, Xin Jin, Dandan Tu, Qingming Huang
HIGHLIGHT: In this work, we propose a new "editing-based" method, i.e., Attribute Group Editing (AGE), for few-shot image generation.

1621, TITLE: Surpassing the Human Accuracy: Detecting Gallbladder Cancer From USG Images With Curriculum Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Basu_Surpassing_the_Human_Accuracy_Detecting_Gallbladder_Cancer_From_USG_Images_CVPR_2022_paper.html
AUTHORS: Soumen Basu, Mayank Gupta, Pratyaksha Rana, Pankaj Gupta, Chetan Arora
HIGHLIGHT: We explore the potential of CNN-based models for gallbladder cancer (GBC) detection from ultrasound (USG) images as no prior study is known.

1622, TITLE: CroMo: Cross-Modal Learning for Monocular Depth Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Verdie_CroMo_Cross-Modal_Learning_for_Monocular_Depth_Estimation_CVPR_2022_paper.html
AUTHORS: Yannick Verdié, Jifei Song, Barnabé Mas, Benjamin Busam, AlesÌ† Leonardis, Steven McDonagh
HIGHLIGHT: In this paper we systematically investigate key trade-offs associated with sensor and modality design choices as well as related model training strategies.

1623, TITLE: Self-Supervised Object Detection From Audio-Visual Correspondence
http://openaccess.thecvf.com//content/CVPR2022/html/Afouras_Self-Supervised_Object_Detection_From_Audio-Visual_Correspondence_CVPR_2022_paper.html
AUTHORS: Triantafyllos Afouras, Yuki M. Asano, Francois Fagan, Andrea Vedaldi, Florian Metze
HIGHLIGHT: We tackle the problem of learning object detectors without supervision.

1624, TITLE: Autofocus for Event Cameras
http://openaccess.thecvf.com//content/CVPR2022/html/Lin_Autofocus_for_Event_Cameras_CVPR_2022_paper.html
AUTHORS: Shijie Lin, Yinqiang Zhang, Lei Yu, Bin Zhou, Xiaowei Luo, Jia Pan
HIGHLIGHT: In particular, the inherent differences between event and frame data in terms of sensing modality, noise, temporal resolutions, etc., bring many challenges in designing an effective AF method for event cameras. To address these challenges, we develop a novel event-based autofocus framework consisting of an event-specific focus measure called event rate (ER) and a robust search strategy called event-based golden search (EGS).

1625, TITLE: Learning Multiple Adverse Weather Removal via Two-Stage Knowledge Learning and Multi-Contrastive Regularization: Toward a Unified Model
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Learning_Multiple_Adverse_Weather_Removal_via_Two-Stage_Knowledge_Learning_and_CVPR_2022_paper.html
AUTHORS: Wei-Ting Chen, Zhi-Kai Huang, Cheng-Che Tsai, Hao-Hsiang Yang, Jian-Jiun Ding, Sy-Yen Kuo
HIGHLIGHT: In this paper, an ill-posed problem of multiple adverse weather removal is investigated.

1626, TITLE: Polymorphic-GAN: Generating Aligned Samples Across Multiple Domains With Learned Morph Maps
http://openaccess.thecvf.com//content/CVPR2022/html/Kim_Polymorphic-GAN_Generating_Aligned_Samples_Across_Multiple_Domains_With_Learned_Morph_CVPR_2022_paper.html
AUTHORS: Seung Wook Kim, Karsten Kreis, Daiqing Li, Antonio Torralba, Sanja Fidler
HIGHLIGHT: In this work, we introduce a generative adversarial network that can simultaneously generate aligned image samples from multiple related domains.

1627, TITLE: Appearance and Structure Aware Robust Deep Visual Graph Matching: Attack, Defense and Beyond
http://openaccess.thecvf.com//content/CVPR2022/html/Ren_Appearance_and_Structure_Aware_Robust_Deep_Visual_Graph_Matching_Attack_CVPR_2022_paper.html
AUTHORS: Qibing Ren, Qingquan Bao, Runzhong Wang, Junchi Yan
HIGHLIGHT: Accordingly, we propose our defense strategy, namely Appearance and Structure Aware Robust Graph Matching (ASAR-GM).

1628, TITLE: Super-Fibonacci Spirals: Fast, Low-Discrepancy Sampling of SO(3)
http://openaccess.thecvf.com//content/CVPR2022/html/Alexa_Super-Fibonacci_Spirals_Fast_Low-Discrepancy_Sampling_of_SO3_CVPR_2022_paper.html
AUTHORS: Marc Alexa
HIGHLIGHT: A comprehensive evaluation comparing to other methods shows that the generated sets of orientations have low discrepancy, minimal spurious components in the power spectrum, and almost identical Voronoi volumes.

1629, TITLE: TrackFormer: Multi-Object Tracking With Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Meinhardt_TrackFormer_Multi-Object_Tracking_With_Transformers_CVPR_2022_paper.html
AUTHORS: Tim Meinhardt, Alexander Kirillov, Laura Leal-Taixé, Christoph Feichtenhofer
HIGHLIGHT: The challenging task of multi-object tracking (MOT) requires simultaneous reasoning about track initialization, identity, and spatio-temporal trajectories. We formulate this task as a frame-to-frame set prediction problem and introduce TrackFormer, an end-to-end trainable MOT approach based on an encoder-decoder Transformer architecture.

1630, TITLE: L-Verse: Bidirectional Generation Between Image and Text
http://openaccess.thecvf.com//content/CVPR2022/html/Kim_L-Verse_Bidirectional_Generation_Between_Image_and_Text_CVPR_2022_paper.html
AUTHORS: Taehoon Kim, Gwangmo Song, Sihaeng Lee, Sangyun Kim, Yewon Seo, Soonyoung Lee, Seung Hwan Kim, Honglak Lee, Kyunghoon Bae
HIGHLIGHT: To better leverage the correlation between image and text, we propose L-Verse, a novel architecture consisting of feature-augmented variational autoencoder (AugVAE) and bidirectional auto-regressive transformer (BiART) for image-to-text and text-to-image generation.

1631, TITLE: PanopticDepth: A Unified Framework for Depth-Aware Panoptic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Gao_PanopticDepth_A_Unified_Framework_for_Depth-Aware_Panoptic_Segmentation_CVPR_2022_paper.html
AUTHORS: Naiyu Gao, Fei He, Jian Jia, Yanhu Shan, Haoyang Zhang, Xin Zhao, Kaiqi Huang
HIGHLIGHT: This paper presents a unified framework for depth-aware panoptic segmentation (DPS), which aims to reconstruct 3D scene with instance-level semantics from one single image.

1632, TITLE: 3D Shape Reconstruction From 2D Images With Disentangled Attribute Flow
http://openaccess.thecvf.com//content/CVPR2022/html/Wen_3D_Shape_Reconstruction_From_2D_Images_With_Disentangled_Attribute_Flow_CVPR_2022_paper.html
AUTHORS: Xin Wen, Junsheng Zhou, Yu-Shen Liu, Hua Su, Zhen Dong, Zhizhong Han
HIGHLIGHT: Since the semantic attributes of a single image are usually implicit and entangled with each other, it is still challenging to reconstruct 3D shape with detailed semantic structures represented by the input image. To address this problem, we propose 3DAttriFlow to disentangle and extract semantic attributes through different semantic levels in the input images.

1633, TITLE: Feature Statistics Mixing Regularization for Generative Adversarial Networks

http://openaccess.thecvf.com//content/CVPR2022/html/Kim_Feature_Statistics_Mixing_Regularization_for_Generative_Adversarial_Networks_CVPR_2022_paper.html
AUTHORS: Junho Kim, Yunjey Choi, Youngjung Uh
HIGHLIGHT: As a remedy, we propose feature statistics mixing regularization (FSMR) that encourages the discriminator's prediction to be invariant to the styles of input images.

1634, TITLE: Learning To Learn and Remember Super Long Multi-Domain Task Sequence
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Learning_To_Learn_and_Remember_Super_Long_Multi-Domain_Task_Sequence_CVPR_2022_paper.html
AUTHORS: Zhenyi Wang, Li Shen, Tiehang Duan, Donglin Zhan, Le Fang, Mingchen Gao
HIGHLIGHT: In this work, we propose a simple yet effective learning to learn approach, i.e., meta optimizer, to mitigate the CF problem in SDML.

1635, TITLE: OpenTAL: Towards Open Set Temporal Action Localization
http://openaccess.thecvf.com//content/CVPR2022/html/Bao_OpenTAL_Towards_Open_Set_Temporal_Action_Localization_CVPR_2022_paper.html
AUTHORS: Wentao Bao, Qi Yu, Yu Kong
HIGHLIGHT: In this paper, we, for the first time, step toward the Open Set TAL (OSTAL) problem and propose a general framework OpenTAL based on Evidential Deep Learning (EDL).

1636, TITLE: Urban Radiance Fields
http://openaccess.thecvf.com//content/CVPR2022/html/Rematas_Urban_Radiance_Fields_CVPR_2022_paper.html
AUTHORS: Konstantinos Rematas, Andrew Liu, Pratul P. Srinivasan, Jonathan T. Barron, Andrea Tagliasacchi, Thomas Funkhouser, Vittorio Ferrari
HIGHLIGHT: The goal of this work is to perform 3D reconstruction and novel view synthesis from data captured by scanning platforms commonly deployed for world mapping in urban outdoor environments (e.g., Street View).

1637, TITLE: Self-Supervised Learning of Adversarial Example: Towards Good Generalizations for Deepfake Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Self-Supervised_Learning_of_Adversarial_Example_Towards_Good_Generalizations_for_Deepfake_CVPR_2022_paper.html
AUTHORS: Liang Chen, Yong Zhang, Yibing Song, Lingqiao Liu, Jue Wang
HIGHLIGHT: This work addresses the generalizable deepfake detection from a simple principle: a generalizable representation should be sensitive to diverse types of forgeries. Following this principle, we propose to enrich the "diversity" of forgeries by synthesizing augmented forgeries with a pool of forgery configurations and strengthen the "sensitivity" to the forgeries by enforcing the model to predict the forgery configurations.

1638, TITLE: Domain-Agnostic Prior for Transfer Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Huo_Domain-Agnostic_Prior_for_Transfer_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Xinyue Huo, Lingxi Xie, Hengtong Hu, Wengang Zhou, Houqiang Li, Qi Tian
HIGHLIGHT: In this paper, we present a simple and effective mechanism that regularizes cross-domain representation learning with a domain-agnostic prior (DAP) that constrains the features extracted from source and target domains to align with a domain-agnostic space.

1639, TITLE: Dynamic Kernel Selection for Improved Generalization and Memory Efficiency in Meta-Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Chavan_Dynamic_Kernel_Selection_for_Improved_Generalization_and_Memory_Efficiency_in_CVPR_2022_paper.html
AUTHORS: Arnav Chavan, Rishabh Tiwari, Udbhav Bamba, Deepak K. Gupta
HIGHLIGHT: In this paper, we present MetaDOCK, a task-specific dynamic kernel selection strategy for designing compressed CNN models that generalize well on unseen tasks in meta-learning.

1640, TITLE: Ego4D: Around the World in 3,000 Hours of Egocentric Video
http://openaccess.thecvf.com//content/CVPR2022/html/Grauman_Ego4D_Around_the_World_in_3000_Hours_of_Egocentric_Video_CVPR_2022_paper.html
AUTHORS: Kristen Grauman, Andrew Westbury, Eugene Byrne, Zachary Chavis, Antonino Furnari, Rohit Girdhar, Jackson Hamburger, Hao Jiang, Miao Liu, Xingyu Liu, Miguel Martin, Tushar Nagarajan, Ilija Radosavovic, Santhosh Kumar Ramakrishnan, Fiona Ryan, Jayant Sharma, Michael Wray, Mengmeng Xu, Eric Zhongcong Xu, Chen Zhao, Siddhant Bansal, Dhruv Batra, Vincent Cartillier, Sean Crane, Tien Do, Morrie Doulaty, Akshay Erapalli, Christoph Feichtenhofer, Adriano Fragomeni, Qichen Fu, Abrham Gebreselasie, Cristina González, James Hillis, Xuhua Huang, Yifei Huang, Wenqi Jia, Weslie Khoo, Jáchym KoláÅ™, Satwik Kottur, Anurag Kumar, Federico Landini, Chao Li, Yanghao Li, Zhenqiang Li, Karttikeya Mangalam, Raghava Modhugu, Jonathan Munro, Tullie Murrell, Takumi Nishiyasu, Will Price, Paola Ruiz, Merey Ramazanova, Leda Sari, Kiran Somasundaram, Audrey Southerland, Yusuke Sugano, Ruijie Tao, Minh Vo, Yuchen Wang, Xindi Wu, Takuma Yagi, Ziwei Zhao, Yunyi Zhu, Pablo Arbeláez, David Crandall, Dima Damen, Giovanni Maria Farinella, Christian Fuegen, Bernard Ghanem, Vamsi Krishna Ithapu, C. V. Jawahar, Hanbyul Joo, Kris Kitani, Haizhou Li, Richard Newcombe, Aude Oliva, Hyun Soo Park, James M. Rehg, Yoichi Sato, Jianbo Shi, Mike Zheng Shou, Antonio Torralba, Lorenzo Torresani, Mingfei Yan, Jitendra Malik
HIGHLIGHT: We introduce Ego4D, a massive-scale egocentric video dataset and benchmark suite.

1641, TITLE: Differentially Private Federated Learning With Local Regularization and Sparsification
http://openaccess.thecvf.com//content/CVPR2022/html/Cheng_Differentially_Private_Federated_Learning_With_Local_Regularization_and_Sparsification_CVPR_2022_paper.html

AUTHORS: Anda Cheng, Peisong Wang, Xi Sheryl Zhang, Jian Cheng
HIGHLIGHT: In this paper, we study the cause of model performance degradation in federated learning with user-level DP guarantee.

1642, TITLE: Self-Supervised Pre-Training of Swin Transformers for 3D Medical Image Analysis
http://openaccess.thecvf.com//content/CVPR2022/html/Tang_Self-Supervised_Pre-Training_of_Swin_Transformers_for_3D_Medical_Image_Analysis_CVPR_2022_paper.html
AUTHORS: Yucheng Tang, Dong Yang, Wenqi Li, Holger R. Roth, Bennett Landman, Daguang Xu, Vishwesh Nath, Ali Hatamizadeh
HIGHLIGHT: Vision Transformers (ViT)s have shown great performance in self-supervised learning of global and local representations that can be transferred to downstream applications. Inspired by these results, we introduce a novel self-supervised learning framework with tailored proxy tasks for medical image analysis.

1643, TITLE: Camera-Conditioned Stable Feature Generation for Isolated Camera Supervised Person Re-IDentification
http://openaccess.thecvf.com//content/CVPR2022/html/Wu_Camera-Conditioned_Stable_Feature_Generation_for_Isolated_Camera_Supervised_Person_Re-IDentification_CVPR_2022_paper.html
AUTHORS: Chao Wu, Wenhang Ge, Ancong Wu, Xiaobin Chang
HIGHLIGHT: However, such cross-view training samples could be unavailable under the ISolated Camera Supervised (ISCS) setting, e.g., a surveillance system deployed across distant scenes. To handle this challenging problem, a new pipeline is introduced by synthesizing the cross-camera samples in the feature space for model training.

1644, TITLE: Weakly Supervised Semantic Segmentation Using Out-of-Distribution Data
http://openaccess.thecvf.com//content/CVPR2022/html/Lee_Weakly_Supervised_Semantic_Segmentation_Using_Out-of-Distribution_Data_CVPR_2022_paper.html
AUTHORS: Jungbeom Lee, Seong Joon Oh, Sangdoo Yun, Junsuk Choe, Eunji Kim, Sungroh Yoon
HIGHLIGHT: We propose a method, W-OoD, for utilizing the hard OoDs.

1645, TITLE: Point-Level Region Contrast for Object Detection Pre-Training
http://openaccess.thecvf.com//content/CVPR2022/html/Bai_Point-Level_Region_Contrast_for_Object_Detection_Pre-Training_CVPR_2022_paper.html
AUTHORS: Yutong Bai, Xinlei Chen, Alexander Kirillov, Alan Yuille, Alexander C. Berg
HIGHLIGHT: In this work we present point-level region contrast, a self-supervised pre-training approach for the task of object detection.

1646, TITLE: Upright-Net: Learning Upright Orientation for 3D Point Cloud
http://openaccess.thecvf.com//content/CVPR2022/html/Pang_Upright-Net_Learning_Upright_Orientation_for_3D_Point_Cloud_CVPR_2022_paper.html
AUTHORS: Xufang Pang, Feng Li, Ning Ding, Xiaopin Zhong
HIGHLIGHT: In this paper, we propose Upright-Net, a deep-learning-based approach for estimating the upright orientation of 3D point clouds.

1647, TITLE: Learning Semantic Associations for Mirror Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Guan_Learning_Semantic_Associations_for_Mirror_Detection_CVPR_2022_paper.html
AUTHORS: Huankang Guan, Jiaying Lin, Rynson W.H. Lau
HIGHLIGHT: We observe that humans tend to place mirrors in relation to certain objects for specific functional purposes, e.g., a mirror above the sink. Inspired by this observation, we propose a model to exploit the semantic associations between the mirror and its surrounding objects for a reliable mirror localization.

1648, TITLE: Spatial-Temporal Parallel Transformer for Arm-Hand Dynamic Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Spatial-Temporal_Parallel_Transformer_for_Arm-Hand_Dynamic_Estimation_CVPR_2022_paper.html
AUTHORS: Shuying Liu, Wenbin Wu, Jiaxian Wu, Yue Lin
HIGHLIGHT: We propose an approach to estimate arm and hand dynamics from monocular video by utilizing the relationship between arm and hand.

1649, TITLE: Failure Modes of Domain Generalization Algorithms
http://openaccess.thecvf.com//content/CVPR2022/html/Galstyan_Failure_Modes_of_Domain_Generalization_Algorithms_CVPR_2022_paper.html
AUTHORS: Tigran Galstyan, Hrayr Harutyunyan, Hrant Khachatrian, Greg Ver Steeg, Aram Galstyan
HIGHLIGHT: In this paper we propose an evaluation framework for domain generalization algorithms that allows decomposition of the error into components capturing distinct aspects of generalization.

1650, TITLE: Geometric and Textural Augmentation for Domain Gap Reduction
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Geometric_and_Textural_Augmentation_for_Domain_Gap_Reduction_CVPR_2022_paper.html
AUTHORS: Xiao-Chang Liu, Yong-Liang Yang, Peter Hall
HIGHLIGHT: These new approaches improve performance, but they ignore the geometric variations in object shape that real art exhibits: artists deform and warp objects for artistic effect. Motivated by this observation, we propose a method to reduce bias by jointly increasing the texture and geometry diversities of the training data.

1651, TITLE: Class Similarity Weighted Knowledge Distillation for Continual Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Phan_Class_Similarity_Weighted_Knowledge_Distillation_for_Continual_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Minh Hieu Phan, The-Anh Ta, Son Lam Phung, Long Tran-Thanh, Abdesselam Bouzerdoum
HIGHLIGHT: We identify a problem in CSS: A model tends to be confused between old and new classes that are visually similar, which makes it forget the old ones. To address this gap, we propose REMINDER - a new CSS framework and a novel class similarity knowledge distillation (CSW-KD) method.

1652, TITLE: DAD-3DHeads: A Large-Scale Dense, Accurate and Diverse Dataset for 3D Head Alignment From a Single Image
http://openaccess.thecvf.com//content/CVPR2022/html/Martyniuk_DAD-3DHeads_A_Large-Scale_Dense_Accurate_and_Diverse_Dataset_for_3D_CVPR_2022_paper.html
AUTHORS: Tetiana Martyniuk, Orest Kupyn, Yana Kurlyak, Igor Krashenyi, Jiří Matas, Viktoriia Sharmanska
HIGHLIGHT: We present DAD-3DHeads, a dense and diverse large-scale dataset, and a robust model for 3D Dense Head Alignment in-the-wild.

1653, TITLE: Reconstructing Surfaces for Sparse Point Clouds With On-Surface Priors
http://openaccess.thecvf.com//content/CVPR2022/html/Ma_Reconstructing_Surfaces_for_Sparse_Point_Clouds_With_On-Surface_Priors_CVPR_2022_paper.html
AUTHORS: Baorui Ma, Yu-Shen Liu, Zhizhong Han
HIGHLIGHT: Our key idea is to infer signed distances by pushing both the query projections to be on the surface and the projection distance to be the minimum.

1654, TITLE: HybridCR: Weakly-Supervised 3D Point Cloud Semantic Segmentation via Hybrid Contrastive Regularization
http://openaccess.thecvf.com//content/CVPR2022/html/Li_HybridCR_Weakly-Supervised_3D_Point_Cloud_Semantic_Segmentation_via_Hybrid_Contrastive_CVPR_2022_paper.html
AUTHORS: Mengtian Li, Yuan Xie, Yunhang Shen, Bo Ke, Ruizhi Qiao, Bo Ren, Shaohui Lin, Lizhuang Ma
HIGHLIGHT: To address the huge labeling cost in large-scale point cloud semantic segmentation, we propose a novel hybrid contrastive regularization (HybridCR) framework in weakly-supervised setting, which obtains competitive performance compared to its fully-supervised counterpart.

1655, TITLE: Fine-Tuning Image Transformers Using Learnable Memory
http://openaccess.thecvf.com//content/CVPR2022/html/Sandler_Fine-Tuning_Image_Transformers_Using_Learnable_Memory_CVPR_2022_paper.html
AUTHORS: Mark Sandler, Andrey Zhmoginov, Max Vladymyrov, Andrew Jackson
HIGHLIGHT: In this paper we propose augmenting Vision Transformer models with learnable memory tokens.

1656, TITLE: Contrastive Conditional Neural Processes
http://openaccess.thecvf.com//content/CVPR2022/html/Ye_Contrastive_Conditional_Neural_Processes_CVPR_2022_paper.html
AUTHORS: Zesheng Ye, Lina Yao
HIGHLIGHT: Instead, noise contrastive estimation might be able to provide more robust representations by learning distributional matching objectives to combat such inherent limitation of generative models. In light of this, we propose to equip CNPs by 1) aligning prediction with encoded ground-truth observation, and 2) decoupling meta-representation adaptation from generative reconstruction.

1657, TITLE: vCLIMB: A Novel Video Class Incremental Learning Benchmark
http://openaccess.thecvf.com//content/CVPR2022/html/Villa_vCLIMB_A_Novel_Video_Class_Incremental_Learning_Benchmark_CVPR_2022_paper.html
AUTHORS: Andrés Villa, Kumail Alhamoud, Victor Escorcia, Fabian Caba, Juan León Alcázar, Bernard Ghanem
HIGHLIGHT: We introduce vCLIMB, a novel video continual learning benchmark.

1658, TITLE: Bending Reality: Distortion-Aware Transformers for Adapting to Panoramic Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Bending_Reality_Distortion-Aware_Transformers_for_Adapting_to_Panoramic_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Jiaming Zhang, Kailun Yang, Chaoxiang Ma, Simon Reiß, Kunyu Peng, Rainer Stiefelhagen
HIGHLIGHT: Distortions and the distinct image-feature distribution in 360deg panoramas impede the transfer from the annotation-rich pinhole domain and therefore come with a big dent in performance. To get around this domain difference and bring together semantic annotations from pinhole- and 360deg surround-visuals, we propose to learn object deformations and panoramic image distortions in the Deformable Patch Embedding (DPE) and Deformable MLP (DMLP) components which blend into our Transformer for PAnoramic Semantic Segmentation (Trans4PASS) model.

1659, TITLE: Sparse and Complete Latent Organization for Geospatial Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Sparse_and_Complete_Latent_Organization_for_Geospatial_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Fengyu Yang, Chenyang Ma
HIGHLIGHT: First, foreground objects are tiny in the remote sensing images and are represented by only a few pixels, which leads to large foreground intra-class variance and undermines the discrimination between foreground classes (issue firstly considered

in this work). Second, background class contains complex context, which results in false alarms due to large background intra-class variance. To alleviate these two issues, we construct a sparse and complete latent structure via prototypes.

1660, TITLE:		Robust Equivariant Imaging: A Fully Unsupervised Framework for Learning To Image From Noisy and Partial Measurements
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Robust_Equivariant_Imaging_A_Fully_Unsupervised_Framework_for_Learning_To_CVPR_2022_paper.html
AUTHORS:		Dongdong Chen, Julián Tachella, Mike E. Davies
HIGHLIGHT:		In this paper, we propose a Robust Equivariant Imaging (REI) framework which can learn to image from noisy partial measurements alone.

1661, TITLE:		Not All Relations Are Equal: Mining Informative Labels for Scene Graph Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Goel_Not_All_Relations_Are_Equal_Mining_Informative_Labels_for_Scene_CVPR_2022_paper.html
AUTHORS:		Arushi Goel, Basura Fernando, Frank Keller, Hakan Bilen
HIGHLIGHT:		Learning on trivial relations that indicate generic spatial configuration like 'on' instead of informative relations such as 'parked on' does not enforce this complex reasoning, harming generalization. To address this problem, we propose a novel framework for SGG training that exploits relation labels based on their informativeness.

1662, TITLE:		Learning To Detect Scene Landmarks for Camera Localization
http://openaccess.thecvf.com//content/CVPR2022/html/Do_Learning_To_Detect_Scene_Landmarks_for_Camera_Localization_CVPR_2022_paper.html
AUTHORS:		Tien Do, Ondrej Miksik, Joseph DeGol, Hyun Soo Park, Sudipta N. Sinha
HIGHLIGHT:		We present a new learned camera localization technique that eliminates the need to store features or a detailed 3D point cloud.

1663, TITLE:		INS-Conv: Incremental Sparse Convolution for Online 3D Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_INS-Conv_Incremental_Sparse_Convolution_for_Online_3D_Segmentation_CVPR_2022_paper.html
AUTHORS:		Leyao Liu, Tian Zheng, Yun-Jou Lin, Kai Ni, Lu Fang
HIGHLIGHT:		We propose INS-Conv, an INcremental Sparse Convolutional network which enables online accurate 3D semantic and instance segmentation.

1664, TITLE:		ST++: Make Self-Training Work Better for Semi-Supervised Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_ST_Make_Self-Training_Work_Better_for_Semi-Supervised_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS:		Lihe Yang, Wei Zhuo, Lei Qi, Yinghuan Shi, Yang Gao
HIGHLIGHT:		Inspired by the impressive results, we thoroughly investigate the SDA and provide some empirical analysis.

1665, TITLE:		Visual Vibration Tomography: Estimating Interior Material Properties From Monocular Video
http://openaccess.thecvf.com//content/CVPR2022/html/Feng_Visual_Vibration_Tomography_Estimating_Interior_Material_Properties_From_Monocular_Video_CVPR_2022_paper.html
AUTHORS:		Berthy T. Feng, Alexander C. Ogren, Chiara Daraio, Katherine L. Bouman
HIGHLIGHT:		An object's interior material properties, while invisible to the human eye, determine motion observed on its surface. We propose an approach that estimates heterogeneous material properties of an object from a monocular video of its surface vibrations.

1666, TITLE:		Self-Supervised Global-Local Structure Modeling for Point Cloud Domain Adaptation With Reliable Voted Pseudo Labels
http://openaccess.thecvf.com//content/CVPR2022/html/Fan_Self-Supervised_Global-Local_Structure_Modeling_for_Point_Cloud_Domain_Adaptation_With_CVPR_2022_paper.html
AUTHORS:		Hehe Fan, Xiaojun Chang, Wanyue Zhang, Yi Cheng, Ying Sun, Mohan Kankanhalli
HIGHLIGHT:		In this paper, we propose an unsupervised domain adaptation method for deep point cloud representation learning.

1667, TITLE:		Interacting Attention Graph for Single Image Two-Hand Reconstruction
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Interacting_Attention_Graph_for_Single_Image_Two-Hand_Reconstruction_CVPR_2022_paper.html
AUTHORS:		Mengcheng Li, Liang An, Hongwen Zhang, Lianpeng Wu, Feng Chen, Tao Yu, Yebin Liu
HIGHLIGHT:		In this paper, we present Interacting Attention Graph Hand (IntagHand), the first graph convolution based network that reconstructs two interacting hands from a single RGB image.

1668, TITLE:		Rope3D: The Roadside Perception Dataset for Autonomous Driving and Monocular 3D Object Detection Task
http://openaccess.thecvf.com//content/CVPR2022/html/Ye_Rope3D_The_Roadside_Perception_Dataset_for_Autonomous_Driving_and_Monocular_CVPR_2022_paper.html
AUTHORS:		Xiaoqing Ye, Mao Shu, Hanyu Li, Yifeng Shi, Yingying Li, Guangjie Wang, Xiao Tan, Errui Ding
HIGHLIGHT:		To accelerate the progress of roadside perception, we present the first high-diversity challenging Roadside Perception 3D dataset- Rope3D from a novel view.

1669, TITLE:		Noisy Boundaries: Lemon or Lemonade for Semi-Supervised Instance Segmentation?

http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Noisy_Boundaries_Lemon_or_Lemonade_for_Semi-Supervised_Instance_Segmentation_CVPR_2022_paper.html
AUTHORS: Zhenyu Wang, Yali Li, Shengjin Wang
HIGHLIGHT: In this paper, we formally address semi-supervised instance segmentation, where unlabeled images are employed to boost the performance.

1670, TITLE: Boosting View Synthesis With Residual Transfer
http://openaccess.thecvf.com//content/CVPR2022/html/Rong_Boosting_View_Synthesis_With_Residual_Transfer_CVPR_2022_paper.html
AUTHORS: Xuejian Rong, Jia-Bin Huang, Ayush Saraf, Changil Kim, Johannes Kopf
HIGHLIGHT: We present a simple but effective technique to boost the rendering quality, which can be easily integrated with most view synthesis methods.

1671, TITLE: Input-Level Inductive Biases for 3D Reconstruction
http://openaccess.thecvf.com//content/CVPR2022/html/Yifan_Input-Level_Inductive_Biases_for_3D_Reconstruction_CVPR_2022_paper.html
AUTHORS: Wang Yifan, Carl Doersch, Relja Arandjelović, João Carreira, Andrew Zisserman
HIGHLIGHT: In this paper we tackle 3D reconstruction using a domain agnostic architecture and study how instead to inject the same type of inductive biases directly as extra inputs to the model.

1672, TITLE: Exploring and Evaluating Image Restoration Potential in Dynamic Scenes
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Exploring_and_Evaluating_Image_Restoration_Potential_in_Dynamic_Scenes_CVPR_2022_paper.html
AUTHORS: Cheng Zhang, Shaolin Su, Yu Zhu, Qingsen Yan, Jinqiu Sun, Yanning Zhang
HIGHLIGHT: In this paper, to better study an image's potential value that can be explored for restoration, we propose a novel concept, referring to image restoration potential (IRP).

1673, TITLE: FashionVLP: Vision Language Transformer for Fashion Retrieval With Feedback
http://openaccess.thecvf.com//content/CVPR2022/html/Goenka_FashionVLP_Vision_Language_Transformer_for_Fashion_Retrieval_With_Feedback_CVPR_2022_paper.html
AUTHORS: Sonam Goenka, Zhaoheng Zheng, Ayush Jaiswal, Rakesh Chada, Yue Wu, Varsha Hedau, Pradeep Natarajan
HIGHLIGHT: We propose a new vision-language transformer based model, FashionVLP, that brings the prior knowledge contained in large image-text corpora to the domain of fashion image re-trieval, and combines visual information from multiple levels of context to effectively capture fashion related information.

1674, TITLE: Cross-Image Relational Knowledge Distillation for Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Cross-Image_Relational_Knowledge_Distillation_for_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Chuanguang Yang, Helong Zhou, Zhulin An, Xue Jiang, Yongjun Xu, Qian Zhang
HIGHLIGHT: This paper proposes a novel Cross-Image Relational KD (CIRKD), which focuses on transferring structured pixel-to-pixel and pixel-to-region relations among the whole images.

1675, TITLE: A-ViT: Adaptive Tokens for Efficient Vision Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Yin_A-ViT_Adaptive_Tokens_for_Efficient_Vision_Transformer_CVPR_2022_paper.html
AUTHORS: Hongxu Yin, Arash Vahdat, Jose M. Alvarez, Arun Mallya, Jan Kautz, Pavlo Molchanov
HIGHLIGHT: We introduce A-ViT, a method that adaptively adjusts the inference cost of vision transformer ViT for images of different complexity.

1676, TITLE: Think Global, Act Local: Dual-Scale Graph Transformer for Vision-and-Language Navigation
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Think_Global_Act_Local_Dual-Scale_Graph_Transformer_for_Vision-and-Language_Navigation_CVPR_2022_paper.html
AUTHORS: Shizhe Chen, Pierre-Louis Guhur, Makarand Tapaswi, Cordelia Schmid, Ivan Laptev
HIGHLIGHT: In this work, we propose a dual-scale graph transformer (DUET) for joint long-term action planning and fine-grained cross-modal understanding.

1677, TITLE: Towards Layer-Wise Image Vectorization
http://openaccess.thecvf.com//content/CVPR2022/html/Ma_Towards_Layer-Wise_Image_Vectorization_CVPR_2022_paper.html
AUTHORS: Xu Ma, Yuqian Zhou, Xingqian Xu, Bin Sun, Valerii Filev, Nikita Orlov, Yun Fu, Humphrey Shi
HIGHLIGHT: In this work, we propose Layer-wise Image Vectorization, namely LIVE, to convert raster images to SVGs and simultaneously maintain its image topology.

1678, TITLE: Scenic: A JAX Library for Computer Vision Research and Beyond
http://openaccess.thecvf.com//content/CVPR2022/html/Dehghani_Scenic_A_JAX_Library_for_Computer_Vision_Research_and_Beyond_CVPR_2022_paper.html
AUTHORS: Mostafa Dehghani, Alexey Gritsenko, Anurag Arnab, Matthias Minderer, Yi Tay
HIGHLIGHT: Scenic is an open-source (https://github.com/google-research/scenic) JAX library with a focus on transformer-based models for computer vision research and beyond.

1679, TITLE: CNN Filter DB: An Empirical Investigation of Trained Convolutional Filters

178

http://openaccess.thecvf.com//content/CVPR2022/html/Gavrikov_CNN_Filter_DB_An_Empirical_Investigation_of_Trained_Convolutional_Filters_CVPR_2022_paper.html
AUTHORS: Paul Gavrikov, Janis Keuper
HIGHLIGHT: While ongoing research efforts are engaging these problems from various angles, in most computer vision related cases these approaches can be generalized to investigations of the effects of distribution shifts in image data. In this context, we propose to study the shifts in the learned weights of trained CNN models.

1680, TITLE: ScePT: Scene-Consistent, Policy-Based Trajectory Predictions for Planning
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_ScePT_Scene-Consistent_Policy-Based_Trajectory_Predictions_for_Planning_CVPR_2022_paper.html
AUTHORS: Yuxiao Chen, Boris Ivanovic, Marco Pavone
HIGHLIGHT: In this work, we present ScePT, a policy planning-based trajectory prediction model that generates accurate, scene-consistent trajectory predictions suitable for autonomous system motion planning.

1681, TITLE: Calibrating Deep Neural Networks by Pairwise Constraints
http://openaccess.thecvf.com//content/CVPR2022/html/Cheng_Calibrating_Deep_Neural_Networks_by_Pairwise_Constraints_CVPR_2022_paper.html
AUTHORS: Jiacheng Cheng, Nuno Vasconcelos
HIGHLIGHT: It is well known that deep neural networks (DNNs) produce poorly calibrated estimates of class-posterior probabilities. We hypothesize that this is due to the limited calibration supervision provided by the cross-entropy loss, which places all emphasis on the probability of the true class and mostly ignores the remaining.

1682, TITLE: Deep Saliency Prior for Reducing Visual Distraction
http://openaccess.thecvf.com//content/CVPR2022/html/Aberman_Deep_Saliency_Prior_for_Reducing_Visual_Distraction_CVPR_2022_paper.html
AUTHORS: Kfir Aberman, Junfeng He, Yossi Gandelsman, Inbar Mosseri, David E. Jacobs, Kai Kohlhoff, Yael Pritch, Michael Rubinstein
HIGHLIGHT: We present results on a variety of natural images and conduct a perceptual study to evaluate and validate the changes in viewers' eye-gaze between the original images and our edited results.

1683, TITLE: Efficient Large-Scale Localization by Global Instance Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Xue_Efficient_Large-Scale_Localization_by_Global_Instance_Recognition_CVPR_2022_paper.html
AUTHORS: Fei Xue, Ignas Budvytis, Daniel Olmeda Reino, Roberto Cipolla
HIGHLIGHT: In this paper, we propose an efficient and accurate large-scale localization framework based on the recognition of buildings, which are not only discriminative for coarse localization but also robust for fine localization.

1684, TITLE: Sign Language Video Retrieval With Free-Form Textual Queries
http://openaccess.thecvf.com//content/CVPR2022/html/Duarte_Sign_Language_Video_Retrieval_With_Free-Form_Textual_Queries_CVPR_2022_paper.html
AUTHORS: Amanda Duarte, Samuel Albanie, Xavier Giró-i-Nieto, Gül Varol
HIGHLIGHT: We identify that a key bottleneck in the performance of the system is the quality of the sign video embedding which suffers from a scarcity of labelled training data. We, therefore, propose SPOT-ALIGN, a framework for interleaving iterative rounds of sign spotting and feature alignment to expand the scope and scale of available training data.

1685, TITLE: Real-Time Object Detection for Streaming Perception
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Real-Time_Object_Detection_for_Streaming_Perception_CVPR_2022_paper.html
AUTHORS: Jinrong Yang, Songtao Liu, Zeming Li, Xiaoping Li, Jian Sun
HIGHLIGHT: While past works ignore the inevitable changes in the environment after processing, streaming perception is proposed to jointly evaluate the latency and accuracy into a single metric for video online perception. In this paper, instead of searching trade-offs between accuracy and speed like previous works, we point out that endowing real-time models with the ability to predict the future is the key to dealing with this problem.

1686, TITLE: Simulated Adversarial Testing of Face Recognition Models
http://openaccess.thecvf.com//content/CVPR2022/html/Ruiz_Simulated_Adversarial_Testing_of_Face_Recognition_Models_CVPR_2022_paper.html
AUTHORS: Nataniel Ruiz, Adam Kortylewski, Weichao Qiu, Cihang Xie, Sarah Adel Bargal, Alan Yuille, Stan Sclaroff
HIGHLIGHT: In this work, we propose a framework for learning how to test machine learning algorithms using simulators in an adversarial manner in order to find weaknesses in the model before deploying it in critical scenarios.

1687, TITLE: VisualHow: Multimodal Problem Solving
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_VisualHow_Multimodal_Problem_Solving_CVPR_2022_paper.html
AUTHORS: Jinhui Yang, Xianyu Chen, Ming Jiang, Shi Chen, Louis Wang, Qi Zhao
HIGHLIGHT: With an overarching goal of developing intelligent systems to assist humans in various daily activities, we propose VisualHow, a free-form and open-ended research that focuses on understanding a real-life problem and deriving its solution by incorporating key components across multiple modalities.

1688, TITLE: Equivariance Allows Handling Multiple Nuisance Variables When Analyzing Pooled Neuroimaging Datasets

http://openaccess.thecvf.com//content/CVPR2022/html/Lokhande_Equivariance_Allows_Handling_Multiple_Nuisance_Variables_When_Analyzing_Pooled_Neuroimaging_CVPR_2022_paper.html
AUTHORS: Vishnu Suresh Lokhande, Rudrasis Chakraborty, Sathya N. Ravi, Vikas Singh
HIGHLIGHT: Invariant representation learning, by itself, is ill-suited to fully model the data generation process. In this paper, we show how bringing recent results on equivariant representation learning (for studying symmetries in neural networks) together with simple use of classical results on causal inference provides an effective practical solution to this problem.

1689, TITLE: Spatial Commonsense Graph for Object Localisation in Partial Scenes
http://openaccess.thecvf.com//content/CVPR2022/html/Giuliari_Spatial_Commonsense_Graph_for_Object_Localisation_in_Partial_Scenes_CVPR_2022_paper.html
AUTHORS: Francesco Giuliari, Geri Skenderi, Marco Cristani, Yiming Wang, Alessio Del Bue
HIGHLIGHT: We solve object localisation in partial scenes, a new problem of estimating the unknown position of an object (e.g. where is the bag?)

1690, TITLE: CAT-Det: Contrastively Augmented Transformer for Multi-Modal 3D Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_CAT-Det_Contrastively_Augmented_Transformer_for_Multi-Modal_3D_Object_Detection_CVPR_2022_paper.html
AUTHORS: Yanan Zhang, Jiaxin Chen, Di Huang
HIGHLIGHT: However, it is quite difficult to sufficiently use them, due to large inter-modal discrepancies. To address this issue, we propose a novel framework, namely Contrastively Augmented Transformer for multi-modal 3D object Detection (CAT-Det).

1691, TITLE: OSSGAN: Open-Set Semi-Supervised Image Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Katsumata_OSSGAN_Open-Set_Semi-Supervised_Image_Generation_CVPR_2022_paper.html
AUTHORS: Kai Katsumata, Duc Minh Vo, Hideki Nakayama
HIGHLIGHT: We introduce a challenging training scheme of conditional GANs, called open-set semi-supervised image generation, where the training dataset consists of two parts: (i) labeled data and (ii) unlabeled data with samples belonging to one of the labeled data classes, namely, a closed-set, and samples not belonging to any of the labeled data classes, namely, an open-set.

1692, TITLE: Lite Vision Transformer With Enhanced Self-Attention
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Lite_Vision_Transformer_With_Enhanced_Self-Attention_CVPR_2022_paper.html
AUTHORS: Chenglin Yang, Yilin Wang, Jianming Zhang, He Zhang, Zijun Wei, Zhe Lin, Alan Yuille
HIGHLIGHT: We propose Lite Vision Transformer (LVT), a novel light-weight transformer network with two enhanced self-attention mechanisms to improve the model performances for mobile deployment.

1693, TITLE: Diversity Matters: Fully Exploiting Depth Clues for Reliable Monocular 3D Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Diversity_Matters_Fully_Exploiting_Depth_Clues_for_Reliable_Monocular_3D_CVPR_2022_paper.html
AUTHORS: Zhuoling Li, Zhan Qu, Yang Zhou, Jianzhuang Liu, Haoqian Wang, Lihui Jiang
HIGHLIGHT: However, these assumptions do not always hold in practical applications. To tackle this problem, we propose a depth solving system that fully explores the visual clues from the subtasks in M3OD and generates multiple estimations for the depth of each target.

1694, TITLE: NinjaDesc: Content-Concealing Visual Descriptors via Adversarial Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Ng_NinjaDesc_Content-Concealing_Visual_Descriptors_via_Adversarial_Learning_CVPR_2022_paper.html
AUTHORS: Tony Ng, Hyo Jin Kim, Vincent T. Lee, Daniel DeTone, Tsun-Yi Yang, Tianwei Shen, Eddy Ilg, Vassileios Balntas, Krystian Mikolajczyk, Chris Sweeney
HIGHLIGHT: In particular, we propose an adversarial learning framework for training visual descriptors that prevent image reconstruction, while maintaining the matching accuracy.

1695, TITLE: Physically-Guided Disentangled Implicit Rendering for 3D Face Modeling
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Physically-Guided_Disentangled_Implicit_Rendering_for_3D_Face_Modeling_CVPR_2022_paper.html
AUTHORS: Zhenyu Zhang, Yanhao Ge, Ying Tai, Weijian Cao, Renwang Chen, Kunlin Liu, Hao Tang, Xiaoming Huang, Chengjie Wang, Zhifeng Xie, Dongjin Huang
HIGHLIGHT: This paper presents a novel Physically-guided Disentangled Implicit Rendering (PhyDIR) framework for high-fidelity 3D face modeling.

1696, TITLE: M5Product: Self-Harmonized Contrastive Learning for E-Commercial Multi-Modal Pretraining
http://openaccess.thecvf.com//content/CVPR2022/html/Dong_M5Product_Self-Harmonized_Contrastive_Learning_for_E-Commercial_Multi-Modal_Pretraining_CVPR_2022_paper.html
AUTHORS: Xiao Dong, Xunlin Zhan, Yangxin Wu, Yunchao Wei, Michael C. Kampffmeyer, Xiaoyong Wei, Minlong Lu, Yaowei Wang, Xiaodan Liang
HIGHLIGHT: By leveraging the natural suitability of E-commerce, where different modalities capture complementary semantic information, we contribute a large-scale multi-modal pre-training dataset M5Product.

1697, TITLE: Bi-Level Alignment for Cross-Domain Crowd Counting

http://openaccess.thecvf.com//content/CVPR2022/html/Gong_Bi-Level_Alignment_for_Cross-Domain_Crowd_Counting_CVPR_2022_paper.html
AUTHORS: Shenjian Gong, Shanshan Zhang, Jian Yang, Dengxin Dai, Bernt Schiele
HIGHLIGHT: In this work, we aim to develop a new adversarial learning based method, which is simple and efficient to apply.

1698, TITLE: ST-MFNet: A Spatio-Temporal Multi-Flow Network for Frame Interpolation
http://openaccess.thecvf.com//content/CVPR2022/html/Danier_ST-MFNet_A_Spatio-Temporal_Multi-Flow_Network_for_Frame_Interpolation_CVPR_2022_paper.html
AUTHORS: Duolikun Danier, Fan Zhang, David Bull
HIGHLIGHT: VFI can be extremely challenging, particularly in sequences containing large motions, occlusions or dynamic textures, where existing approaches fail to offer perceptually robust interpolation performance. In this context, we present a novel deep learning based VFI method, ST-MFNet, based on a Spatio-Temporal Multi-Flow architecture.

1699, TITLE: Self-Supervised Super-Resolution for Multi-Exposure Push-Frame Satellites
http://openaccess.thecvf.com//content/CVPR2022/html/Nguyen_Self-Supervised_Super-Resolution_for_Multi-Exposure_Push-Frame_Satellites_CVPR_2022_paper.html
AUTHORS: Ngoc Long Nguyen, Jérémy Anger, Axel Davy, Pablo Arias, Gabriele Facciolo
HIGHLIGHT: Modern Earth observation satellites capture multi-exposure bursts of push-frame images that can be super-resolved via computational means. In this work, we propose a super-resolution method for such multi-exposure sequences, a problem that has received very little attention in the literature.

1700, TITLE: Efficient Multi-View Stereo by Iterative Dynamic Cost Volume
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Efficient_Multi-View_Stereo_by_Iterative_Dynamic_Cost_Volume_CVPR_2022_paper.html
AUTHORS: Shaoqian Wang, Bo Li, Yuchao Dai
HIGHLIGHT: In this paper, we propose a novel iterative dynamic cost volume for multi-view stereo.

1701, TITLE: Learning To Generate Line Drawings That Convey Geometry and Semantics
http://openaccess.thecvf.com//content/CVPR2022/html/Chan_Learning_To_Generate_Line_Drawings_That_Convey_Geometry_and_Semantics_CVPR_2022_paper.html
AUTHORS: Caroline Chan, Frédo Durand, Phillip Isola
HIGHLIGHT: This paper presents an unpaired method for creating line drawings from photographs.

1702, TITLE: On Guiding Visual Attention With Language Specification
http://openaccess.thecvf.com//content/CVPR2022/html/Petryk_On_Guiding_Visual_Attention_With_Language_Specification_CVPR_2022_paper.html
AUTHORS: Suzanne Petryk, Lisa Dunlap, Keyan Nasseri, Joseph Gonzalez, Trevor Darrell, Anna Rohrbach
HIGHLIGHT: Our insight is to use high-level language specification as advice for constraining the prediction evidence to task-relevant features, instead of distractors.

1703, TITLE: ReSTR: Convolution-Free Referring Image Segmentation Using Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Kim_ReSTR_Convolution-Free_Referring_Image_Segmentation_Using_Transformers_CVPR_2022_paper.html
AUTHORS: Namyup Kim, Dongwon Kim, Cuiling Lan, Wenjun Zeng, Suha Kwak
HIGHLIGHT: Most of existing methods for this task rely heavily on convolutional neural networks, which however have trouble capturing long-range dependencies between entities in the language expression and are not flexible enough for modeling interactions between the two different modalities. To address these issues, we present the first convolution-free model for referring image segmentation using transformers, dubbed ReSTR.

1704, TITLE: A Graph Matching Perspective With Transformers on Video Instance Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Qin_A_Graph_Matching_Perspective_With_Transformers_on_Video_Instance_Segmentation_CVPR_2022_paper.html
AUTHORS: Zheyun Qin, Xiankai Lu, Xiushan Nie, Yilong Yin, Jianbing Shen
HIGHLIGHT: This paper presents a graph matching-based method to formulate VIS.

1705, TITLE: TransEditor: Transformer-Based Dual-Space GAN for Highly Controllable Facial Editing
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_TransEditor_Transformer-Based_Dual-Space_GAN_for_Highly_Controllable_Facial_Editing_CVPR_2022_paper.html
AUTHORS: Yanbo Xu, Yueqin Yin, Liming Jiang, Qianyi Wu, Chengyao Zheng, Chen Change Loy, Bo Dai, Wayne Wu
HIGHLIGHT: In this study, we highlight the importance of interaction in a dual-space GAN for more controllable editing.

1706, TITLE: FLAG: Flow-Based 3D Avatar Generation From Sparse Observations
http://openaccess.thecvf.com//content/CVPR2022/html/Aliakbarian_FLAG_Flow-Based_3D_Avatar_Generation_From_Sparse_Observations_CVPR_2022_paper.html
AUTHORS: Sadegh Aliakbarian, Pashmina Cameron, Federica Bogo, Andrew Fitzgibbon, Thomas J. Cashman
HIGHLIGHT: While these signals are valuable, they are an incomplete representation of the human body, making it challenging to generate a faithful full-body avatar. We address this challenge by developing a flow-based generative model of the 3D human body from sparse observations, wherein we learn not only a conditional distribution of 3D human pose, but also a probabilistic

mapping from observations to the latent space from which we can generate a plausible pose along with uncertainty estimates for the joints.

1707, TITLE: Stability-Driven Contact Reconstruction From Monocular Color Images
http://openaccess.thecvf.com//content/CVPR2022/html/Zhao_Stability-Driven_Contact_Reconstruction_From_Monocular_Color_Images_CVPR_2022_paper.html
AUTHORS: Zimeng Zhao, Binghui Zuo, Wei Xie, Yangang Wang
HIGHLIGHT: Our key idea is to reconstruct the contact pattern directly from monocular images and utilize the physical stability criterion in the simulation to drive the optimization process described above.

1708, TITLE: Use All the Labels: A Hierarchical Multi-Label Contrastive Learning Framework
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Use_All_the_Labels_A_Hierarchical_Multi-Label_Contrastive_Learning_Framework_CVPR_2022_paper.html
AUTHORS: Shu Zhang, Ran Xu, Caiming Xiong, Chetan Ramaiah
HIGHLIGHT: In this paper, we present a hierarchical multi-label representation learning framework that can leverage all available labels and preserve the hierarchical relationship between classes.

1709, TITLE: SGTR: End-to-End Scene Graph Generation With Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Li_SGTR_End-to-End_Scene_Graph_Generation_With_Transformer_CVPR_2022_paper.html
AUTHORS: Rongjie Li, Songyang Zhang, Xuming He
HIGHLIGHT: In this work, we propose a novel SGG method to address the aforementioned issues, formulating the task as a bipartite graph construction problem.

1710, TITLE: Panoptic Neural Fields: A Semantic Object-Aware Neural Scene Representation
http://openaccess.thecvf.com//content/CVPR2022/html/Kundu_Panoptic_Neural_Fields_A_Semantic_Object-Aware_Neural_Scene_Representation_CVPR_2022_paper.html
AUTHORS: Abhijit Kundu, Kyle Genova, Xiaoqi Yin, Alireza Fathi, Caroline Pantofaru, Leonidas J. Guibas, Andrea Tagliasacchi, Frank Dellaert, Thomas Funkhouser
HIGHLIGHT: We present PanopticNeRF, an object-aware neural scene representation that decomposes a scene into a set of objects (things) and background (stuff).

1711, TITLE: Texture-Based Error Analysis for Image Super-Resolution
http://openaccess.thecvf.com//content/CVPR2022/html/Magid_Texture-Based_Error_Analysis_for_Image_Super-Resolution_CVPR_2022_paper.html
AUTHORS: Salma Abdel Magid, Zudi Lin, Donglai Wei, Yulun Zhang, Jinjin Gu, Hanspeter Pfister
HIGHLIGHT: Our key contribution is to leverage a texture classifier, which enables us to assign patches with semantic labels, to identify the source of SR errors both globally and locally.

1712, TITLE: PILC: Practical Image Lossless Compression With an End-to-End GPU Oriented Neural Framework
http://openaccess.thecvf.com//content/CVPR2022/html/Kang_PILC_Practical_Image_Lossless_Compression_With_an_End-to-End_GPU_Oriented_CVPR_2022_paper.html
AUTHORS: Ning Kang, Shanzhao Qiu, Shifeng Zhang, Zhenguo Li, Shu-Tao Xia
HIGHLIGHT: In this paper, we propose PILC, an end-to-end image lossless compression framework that achieves 200 MB/s for both compression and decompression with a single NVIDIA Tesla V100 GPU, 10x faster than the most efficient one before.

1713, TITLE: Set-Supervised Action Learning in Procedural Task Videos via Pairwise Order Consistency
http://openaccess.thecvf.com//content/CVPR2022/html/Lu_Set-Supervised_Action_Learning_in_Procedural_Task_Videos_via_Pairwise_Order_CVPR_2022_paper.html
AUTHORS: Zijia Lu, Ehsan Elhamifar
HIGHLIGHT: Our key observation is that videos within the same task have similar ordering of actions, which can be leveraged for effective learning. Therefore, we propose an attention-based method with a new Pairwise Ordering Consistency (POC) loss that encourages that for each common action pair in two videos of the same task, the attentions of actions follow a similar ordering.

1714, TITLE: Learning To Align Sequential Actions in the Wild
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Learning_To_Align_Sequential_Actions_in_the_Wild_CVPR_2022_paper.html
AUTHORS: Weizhe Liu, Bugra Tekin, Huseyin Coskun, Vibhav Vineet, Pascal Fua, Marc Pollefeys
HIGHLIGHT: In this paper, we propose an approach to align sequential actions in the wild that involve diverse temporal variations.

1715, TITLE: Decoupled Knowledge Distillation
http://openaccess.thecvf.com//content/CVPR2022/html/Zhao_Decoupled_Knowledge_Distillation_CVPR_2022_paper.html
AUTHORS: Borui Zhao, Quan Cui, Renjie Song, Yiyu Qiu, Jiajun Liang
HIGHLIGHT: More importantly, we reveal that the classical KD loss is a coupled formulation, which (1) suppresses the effectiveness of NCKD and (2) limits the flexibility to balance these two parts. To address these issues, we present Decoupled Knowledge Distillation(DKD), enabling TCKD and NCKD to play their roles more efficiently and flexibly.

1716, TITLE: DeepFusion: Lidar-Camera Deep Fusion for Multi-Modal 3D Object Detection

http://openaccess.thecvf.com//content/CVPR2022/html/Li_DeepFusion_Lidar-Camera_Deep_Fusion_for_Multi-Modal_3D_Object_Detection_CVPR_2022_paper.html
AUTHORS: Yingwei Li, Adams Wei Yu, Tianjian Meng, Ben Caine, Jiquan Ngiam, Daiyi Peng, Junyang Shen, Yifeng Lu, Denny Zhou, Quoc V. Le, Alan Yuille, Mingxing Tan
HIGHLIGHT: In this paper, we propose two novel techniques: InverseAug that inverses geometric-related augmentations, e.g., rotation, to enable accurate geometric alignment between lidar points and image pixels, and LearnableAlign that leverages cross-attention to dynamically capture the correlations between image and lidar features during fusion.

1717, TITLE: Neural Volumetric Object Selection
http://openaccess.thecvf.com//content/CVPR2022/html/Ren_Neural_Volumetric_Object_Selection_CVPR_2022_paper.html
AUTHORS: Zhongzheng Ren, Aseem Agarwala, Bryan Russell, Alexander G. Schwing, Oliver Wang
HIGHLIGHT: We introduce an approach for selecting objects in neural volumetric 3D representations, such as multi-plane images (MPI) and neural radiance fields (NeRF).

1718, TITLE: GCR: Gradient Coreset Based Replay Buffer Selection for Continual Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Tiwari_GCR_Gradient_Coreset_Based_Replay_Buffer_Selection_for_Continual_Learning_CVPR_2022_paper.html
AUTHORS: Rishabh Tiwari, Krishnateja Killamsetty, Rishabh Iyer, Pradeep Shenoy
HIGHLIGHT: We propose Gradient Coreset Replay (GCR), a novel strategy for replay buffer selection and update using a carefully designed optimization criterion.

1719, TITLE: PointCLIP: Point Cloud Understanding by CLIP
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_PointCLIP_Point_Cloud_Understanding_by_CLIP_CVPR_2022_paper.html
AUTHORS: Renrui Zhang, Ziyu Guo, Wei Zhang, Kunchang Li, Xupeng Miao, Bin Cui, Yu Qiao, Peng Gao, Hongsheng Li
HIGHLIGHT: However, it remains under explored that whether CLIP, pre-trained by large-scale image-text pairs in 2D, can be generalized to 3D recognition. In this paper, we identify such a setting is feasible by proposing PointCLIP, which conducts alignment between CLIP-encoded point cloud and 3D category texts.

1720, TITLE: NeRFusion: Fusing Radiance Fields for Large-Scale Scene Reconstruction
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_NeRFusion_Fusing_Radiance_Fields_for_Large-Scale_Scene_Reconstruction_CVPR_2022_paper.html
AUTHORS: Xiaoshuai Zhang, Sai Bi, Kalyan Sunkavalli, Hao Su, Zexiang Xu
HIGHLIGHT: We propose NeRFusion, a method that combines the advantages of NeRF and TSDF-based fusion techniques to achieve efficient large-scale reconstruction and photo-realistic rendering.

1721, TITLE: DeepFace-EMD: Re-Ranking Using Patch-Wise Earth Mover's Distance Improves Out-of-Distribution Face Identification
http://openaccess.thecvf.com//content/CVPR2022/html/Phan_DeepFace-EMD_Re-Ranking_Using_Patch-Wise_Earth_Movers_Distance_Improves_Out-of-Distribution_Face_CVPR_2022_paper.html
AUTHORS: Hai Phan, Anh Nguyen
HIGHLIGHT: Here, we propose a re-ranking approach that compares two faces using the Earth Mover's Distance on the deep, spatial features of image patches.

1722, TITLE: A Sampling-Based Approach for Efficient Clustering in Large Datasets
http://openaccess.thecvf.com//content/CVPR2022/html/Exarchakis_A_Sampling-Based_Approach_for_Efficient_Clustering_in_Large_Datasets_CVPR_2022_paper.html
AUTHORS: Georgios Exarchakis, Omar Oubari, Gregor Lenz
HIGHLIGHT: We propose a simple and efficient clustering method for high-dimensional data with a large number of clusters.

1723, TITLE: General Facial Representation Learning in a Visual-Linguistic Manner
http://openaccess.thecvf.com//content/CVPR2022/html/Zheng_General_Facial_Representation_Learning_in_a_Visual-Linguistic_Manner_CVPR_2022_paper.html
AUTHORS: Yinglin Zheng, Hao Yang, Ting Zhang, Jianmin Bao, Dongdong Chen, Yangyu Huang, Lu Yuan, Dong Chen, Ming Zeng, Fang Wen
HIGHLIGHT: In this paper, we study the transfer performance of pre-trained models on face analysis tasks and introduce a framework, called FaRL, for general facial representation learning.

1724, TITLE: Deep Color Consistent Network for Low-Light Image Enhancement
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Deep_Color_Consistent_Network_for_Low-Light_Image_Enhancement_CVPR_2022_paper.html
AUTHORS: Zhao Zhang, Huan Zheng, Richang Hong, Mingliang Xu, Shuicheng Yan, Meng Wang
HIGHLIGHT: In this paper, we decouple a color image to two main components, a gray image and a color hist histogram.

1725, TITLE: AdaSTE: An Adaptive Straight-Through Estimator To Train Binary Neural Networks
http://openaccess.thecvf.com//content/CVPR2022/html/Le_AdaSTE_An_Adaptive_Straight-Through_Estimator_To_Train_Binary_Neural_Networks_CVPR_2022_paper.html
AUTHORS: Huu Le, Rasmus Kjær Høier, Che-Tsung Lin, Christopher Zach
HIGHLIGHT: We propose a new algorithm for training deep neural networks (DNNs) with binary weights.

1726, TITLE: Reusing the Task-Specific Classifier as a Discriminator: Discriminator-Free Adversarial Domain Adaptation
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Reusing_the_Task-Specific_Classifier_as_a_Discriminator_Discriminator-Free_Adversarial_Domain_CVPR_2022_paper.html
AUTHORS: Lin Chen, Huaian Chen, Zhixiang Wei, Xin Jin, Xiao Tan, Yi Jin, Enhong Chen
HIGHLIGHT: However, most of these methods failed to effectively leverage the predicted discriminative information, and thus cause mode collapse for generator. In this work, we address this problem from a different perspective and design a simple yet effective adversarial paradigm in the form of a discriminator-free adversarial learning network (DALN), wherein the category classifier is reused as a discriminator, which achieves explicit domain alignment and category distinguishment through a unified objective, enabling the DALN to leverage the predicted discriminative information for sufficient feature alignment.

1727, TITLE: Pooling Revisited: Your Receptive Field Is Suboptimal
http://openaccess.thecvf.com//content/CVPR2022/html/Jang_Pooling_Revisited_Your_Receptive_Field_Is_Suboptimal_CVPR_2022_paper.html
AUTHORS: Dong-Hwan Jang, Sanghyeok Chu, Joonhyuk Kim, Bohyung Han
HIGHLIGHT: Hence, we propose a simple yet effective Dynamically Optimized Pooling operation, referred to as DynOPool, which learns the optimized scale factors of feature maps end-to-end.

1728, TITLE: Dual Task Learning by Leveraging Both Dense Correspondence and Mis-Correspondence for Robust Change Detection With Imperfect Matches
http://openaccess.thecvf.com//content/CVPR2022/html/Park_Dual_Task_Learning_by_Leveraging_Both_Dense_Correspondence_and_Mis-Correspondence_CVPR_2022_paper.html
AUTHORS: Jin-Man Park, Ue-Hwan Kim, Seon-Hoon Lee, Jong-Hwan Kim
HIGHLIGHT: Thus, contemporary approaches fail to show the reported performance in real-world settings. To overcome this limitation, we propose SimSaC.

1729, TITLE: Show Me What and Tell Me How: Video Synthesis via Multimodal Conditioning
http://openaccess.thecvf.com//content/CVPR2022/html/Han_Show_Me_What_and_Tell_Me_How_Video_Synthesis_via_CVPR_2022_paper.html
AUTHORS: Ligong Han, Jian Ren, Hsin-Ying Lee, Francesco Barbieri, Kyle Olszewski, Shervin Minaee, Dimitris Metaxas, Sergey Tulyakov
HIGHLIGHT: This work presents a multimodal video generation framework that benefits from text and images provided jointly or separately.

1730, TITLE: Patch Slimming for Efficient Vision Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Tang_Patch_Slimming_for_Efficient_Vision_Transformers_CVPR_2022_paper.html
AUTHORS: Yehui Tang, Kai Han, Yunhe Wang, Chang Xu, Jianyuan Guo, Chao Xu, Dacheng Tao
HIGHLIGHT: Considering that the attention mechanism aggregates different patches layer-by-layer, we present a novel patch slimming approach that discards useless patches in a top-down paradigm.

1731, TITLE: Bijective Mapping Network for Shadow Removal
http://openaccess.thecvf.com//content/CVPR2022/html/Zhu_Bijective_Mapping_Network_for_Shadow_Removal_CVPR_2022_paper.html
AUTHORS: Yurui Zhu, Jie Huang, Xueyang Fu, Feng Zhao, Qibin Sun, Zheng-Jun Zha
HIGHLIGHT: In this work, we argue that shadow removal and generation are interrelated and could provide useful informative supervision for each other.

1732, TITLE: End-to-End Semi-Supervised Learning for Video Action Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Kumar_End-to-End_Semi-Supervised_Learning_for_Video_Action_Detection_CVPR_2022_paper.html
AUTHORS: Akash Kumar, Yogesh Singh Rawat
HIGHLIGHT: In this work, we focus on semi-supervised learning for video action detection which utilizes both labeled as well as unlabeled data.

1733, TITLE: Causal Transportability for Visual Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Mao_Causal_Transportability_for_Visual_Recognition_CVPR_2022_paper.html
AUTHORS: Chengzhi Mao, Kevin Xia, James Wang, Hao Wang, Junfeng Yang, Elias Bareinboim, Carl Vondrick
HIGHLIGHT: However, we then show that the causal effect, which severs all sources of confounding, remains invariant across domains. This motivates us to develop an algorithm to estimate the causal effect for image classification, which is transportable (i.e., invariant) across source and target environments.

1734, TITLE: Local Attention Pyramid for Scene Image Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Shim_Local_Attention_Pyramid_for_Scene_Image_Generation_CVPR_2022_paper.html
AUTHORS: Sang-Heon Shim, Sangeek Hyun, DaeHyun Bae, Jae-Pil Heo
HIGHLIGHT: Specifically, the synthesized qualities of small and less frequent object classes tend to be low. To address this, we propose a novel attention module, Local Attention Pyramid (LAP) module tailored for scene image synthesis, that encourages GANs to generate diverse object classes in a high quality by explicit spread of high attention scores to local regions, since objects in scene images are scattered over the entire images.

1735, TITLE: Multi-Objective Diverse Human Motion Prediction With Knowledge Distillation
http://openaccess.thecvf.com//content/CVPR2022/html/Ma_Multi-Objective_Diverse_Human_Motion_Prediction_With_Knowledge_Distillation_CVPR_2022_paper.html
AUTHORS: Hengbo Ma, Jiachen Li, Ramtin Hosseini, Masayoshi Tomizuka, Chiho Choi
HIGHLIGHT: In this work, we aim to design a prediction framework that can balance the accuracy sampling and diversity sampling during the testing phase.

1736, TITLE: GridShift: A Faster Mode-Seeking Algorithm for Image Segmentation and Object Tracking
http://openaccess.thecvf.com//content/CVPR2022/html/Kumar_GridShift_A_Faster_Mode-Seeking_Algorithm_for_Image_Segmentation_and_Object_CVPR_2022_paper.html
AUTHORS: Abhishek Kumar, Oladayo S. Ajani, Swagatam Das, Rammohan Mallipeddi
HIGHLIGHT: Therefore, it is very slow for large-scale datasets. To address this issue, we propose a mode-seeking algorithm, GridShift, with faster computing and principally based on MeanShift that uses a grid-based approach.

1737, TITLE: Confidence Propagation Cluster: Unleash Full Potential of Object Detectors
http://openaccess.thecvf.com//content/CVPR2022/html/Shen_Confidence_Propagation_Cluster_Unleash_Full_Potential_of_Object_Detectors_CVPR_2022_paper.html
AUTHORS: Yichun Shen, Wanli Jiang, Zhen Xu, Rundong Li, Junghyun Kwon, Siyi Li
HIGHLIGHT: In this paper, inspired by belief propagation (BP), we propose the Confidence Propagation Cluster (CP-Cluster) to replace NMS-based methods, which is fully parallelizable as well as better in accuracy.

1738, TITLE: Cluster-Guided Image Synthesis With Unconditional Models
http://openaccess.thecvf.com//content/CVPR2022/html/Georgopoulos_Cluster-Guided_Image_Synthesis_With_Unconditional_Models_CVPR_2022_paper.html
AUTHORS: Markos Georgopoulos, James Oldfield, Grigorios G. Chrysos, Yannis Panagakis
HIGHLIGHT: In this work, we focus on controllable image generation by leveraging GANs that are well-trained in an unsupervised fashion.

1739, TITLE: ISNet: Shape Matters for Infrared Small Target Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_ISNet_Shape_Matters_for_Infrared_Small_Target_Detection_CVPR_2022_paper.html
AUTHORS: Mingjin Zhang, Rui Zhang, Yuxiang Yang, Haichen Bai, Jing Zhang, Jie Guo
HIGHLIGHT: How to detect the precise shape information of infrared targets remains challenging. In this paper, we propose a novel infrared shape network (ISNet), where Taylor finite difference (TFD)-inspired edge block and two-orientation attention aggregation (TOAA) block are devised to address this problem.

1740, TITLE: Robust Region Feature Synthesizer for Zero-Shot Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Huang_Robust_Region_Feature_Synthesizer_for_Zero-Shot_Object_Detection_CVPR_2022_paper.html
AUTHORS: Peiliang Huang, Junwei Han, De Cheng, Dingwen Zhang
HIGHLIGHT: Zero-shot object detection aims at incorporating class semantic vectors to realize the detection of (both seen and) unseen classes given an unconstrained test image. In this study, we reveal the core challenges in this research area: how to synthesize robust region features (for unseen objects) that are as intra-class diverse and inter-class separable as the real samples, so that strong unseen object detectors can be trained upon them.

1741, TITLE: Virtual Correspondence: Humans as a Cue for Extreme-View Geometry
http://openaccess.thecvf.com//content/CVPR2022/html/Ma_Virtual_Correspondence_Humans_as_a_Cue_for_Extreme-View_Geometry_CVPR_2022_paper.html
AUTHORS: Wei-Chiu Ma, Anqi Joyce Yang, Shenlong Wang, Raquel Urtasun, Antonio Torralba
HIGHLIGHT: We introduce a method to find virtual correspondences based on humans in the scene.

1742, TITLE: Segment, Magnify and Reiterate: Detecting Camouflaged Objects the Hard Way
http://openaccess.thecvf.com//content/CVPR2022/html/Jia_Segment_Magnify_and_Reiterate_Detecting_Camouflaged_Objects_the_Hard_Way_CVPR_2022_paper.html
AUTHORS: Qi Jia, Shuilian Yao, Yu Liu, Xin Fan, Risheng Liu, Zhongxuan Luo
HIGHLIGHT: To tackle camouflaged object detection (COD), we are inspired by humans attention coupled with the coarse-to-fine detection strategy, and thereby propose an iterative refinement framework, coined SegMaR, which integrates Segment, Magnify and Reiterate in a multi-stage detection fashion.

1743, TITLE: SIMBAR: Single Image-Based Scene Relighting for Effective Data Augmentation for Automated Driving Vision Tasks
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_SIMBAR_Single_Image-Based_Scene_Relighting_for_Effective_Data_Augmentation_for_CVPR_2022_paper.html
AUTHORS: Xianling Zhang, Nathan Tseng, Ameerah Syed, Rohan Bhasin, Nikita Jaipuria
HIGHLIGHT: This paper presents a novel image-based relighting pipeline, SIMBAR, that can work with a single image as input.

1744, TITLE: Shape From Thermal Radiation: Passive Ranging Using Multi-Spectral LWIR Measurements

http://openaccess.thecvf.com//content/CVPR2022/html/Nagase_Shape_From_Thermal_Radiation_Passive_Ranging_Using_Multi-Spectral_LWIR_Measurements_CVPR_2022_paper.html
AUTHORS: Yasuto Nagase, Takahiro Kushida, Kenichiro Tanaka, Takuya Funatomi, Yasuhiro Mukaigawa
HIGHLIGHT: In this paper, we propose a new cue of depth sensing using thermal radiation.

1745, TITLE: Multi-Label Classification With Partial Annotations Using Class-Aware Selective Loss
http://openaccess.thecvf.com//content/CVPR2022/html/Ben-Baruch_Multi-Label_Classification_With_Partial_Annotations_Using_Class-Aware_Selective_Loss_CVPR_2022_paper.html
AUTHORS: Emanuel Ben-Baruch, Tal Ridnik, Itamar Friedman, Avi Ben-Cohen, Nadav Zamir, Asaf Noy, Lihi Zelnik-Manor
HIGHLIGHT: We propose to estimate the class distribution using a dedicated temporary model, and we show its improved efficiency over a naive estimation computed using the dataset's partial annotations.

1746, TITLE: HSC4D: Human-Centered 4D Scene Capture in Large-Scale Indoor-Outdoor Space Using Wearable IMUs and LiDAR
http://openaccess.thecvf.com//content/CVPR2022/html/Dai_HSC4D_Human-Centered_4D_Scene_Capture_in_Large-Scale_Indoor-Outdoor_Space_Using_CVPR_2022_paper.html
AUTHORS: Yudi Dai, Yitai Lin, Chenglu Wen, Siqi Shen, Lan Xu, Jingyi Yu, Yuexin Ma, Cheng Wang
HIGHLIGHT: We propose Human-centered 4D Scene Capture (HSC4D) to accurately and efficiently create a dynamic digital world, containing large-scale indoor-outdoor scenes, diverse human motions, and rich interactions between humans and environments.

1747, TITLE: CADTransformer: Panoptic Symbol Spotting Transformer for CAD Drawings
http://openaccess.thecvf.com//content/CVPR2022/html/Fan_CADTransformer_Panoptic_Symbol_Spotting_Transformer_for_CAD_Drawings_CVPR_2022_paper.html
AUTHORS: Zhiwen Fan, Tianlong Chen, Peihao Wang, Zhangyang Wang
HIGHLIGHT: Existing methods, based on convolutional neural networks (CNNs) and/or graph neural networks (GNNs), regress instance bounding boxes in the pixel domain and then convert the predictions into symbols. In this paper, we present a novel framework named CADTransformer, that can painlessly modify existing vision transformer (ViT) backbones to tackle the above limitations for the panoptic symbol spotting task.

1748, TITLE: IntraQ: Learning Synthetic Images With Intra-Class Heterogeneity for Zero-Shot Network Quantization
http://openaccess.thecvf.com//content/CVPR2022/html/Zhong_IntraQ_Learning_Synthetic_Images_With_Intra-Class_Heterogeneity_for_Zero-Shot_Network_CVPR_2022_paper.html
AUTHORS: Yunshan Zhong, Mingbao Lin, Gongrui Nan, Jianzhuang Liu, Baochang Zhang, Yonghong Tian, Rongrong Ji
HIGHLIGHT: Learning to synthesize data has emerged as a promising direction in zero-shot quantization (ZSQ), which represents neural networks by low-bit integer without accessing any of the real data. In this paper, we observe an interesting phenomenon of intra-class heterogeneity in real data and show that existing methods fail to retain this property in their synthetic images, which causes a limited performance increase.

1749, TITLE: M3L: Language-Based Video Editing via Multi-Modal Multi-Level Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Fu_M3L_Language-Based_Video_Editing_via_Multi-Modal_Multi-Level_Transformers_CVPR_2022_paper.html
AUTHORS: Tsu-Jui Fu, Xin Eric Wang, Scott T. Grafton, Miguel P. Eckstein, William Yang Wang
HIGHLIGHT: This paper introduces the language-based video editing (LBVE) task, which allows the model to edit, guided by text instruction, a source video into a target video.

1750, TITLE: I M Avatar: Implicit Morphable Head Avatars From Videos
http://openaccess.thecvf.com//content/CVPR2022/html/Zheng_I_M_Avatar_Implicit_Morphable_Head_Avatars_From_Videos_CVPR_2022_paper.html
AUTHORS: Yufeng Zheng, Victoria Fernández Abrevaya, Marcel C. Bühler, Xu Chen, Michael J. Black, Otmar Hilliges
HIGHLIGHT: Neural volumetric representations approach photorealism but are hard to animate and do not generalize well to unseen expressions. To tackle this problem, we propose IMavatar (Implicit Morphable avatar), a novel method for learning implicit head avatars from monocular videos.

1751, TITLE: BodyMap: Learning Full-Body Dense Correspondence Map
http://openaccess.thecvf.com//content/CVPR2022/html/Ianina_BodyMap_Learning_Full-Body_Dense_Correspondence_Map_CVPR_2022_paper.html
AUTHORS: Anastasia Ianina, Nikolaos Sarafianos, Yuanlu Xu, Ignacio Rocco, Tony Tung
HIGHLIGHT: In this paper we present BodyMap, a new framework for obtaining high-definition full-body and continuous dense correspondence between in-the-wild images of clothed humans and the surface of a 3D template model.

1752, TITLE: Weakly-Supervised Metric Learning With Cross-Module Communications for the Classification of Anterior Chamber Angle Images
http://openaccess.thecvf.com//content/CVPR2022/html/Huang_Weakly-Supervised_Metric_Learning_With_Cross-Module_Communications_for_the_Classification_of_CVPR_2022_paper.html
AUTHORS: Jingqi Huang, Yue Ning, Dong Nie, Linan Guan, Xiping Jia
HIGHLIGHT: In this paper, we propose a novel end-to-end framework GCNet for automated Glaucoma Classification based on ACA images or other Glaucoma-related medical images.

1753, TITLE: A Hybrid Egocentric Activity Anticipation Framework via Memory-Augmented Recurrent and One-Shot Representation Forecasting
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_A_Hybrid_Egocentric_Activity_Anticipation_Framework_via_Memory-Augmented_Recurrent_and_CVPR_2022_paper.html
AUTHORS: Tianshan Liu, Kin-Man Lam
HIGHLIGHT: However, the limitations of current recursive prediction models arise from two aspects: (i) The vanilla recurrent units are prone to accumulated errors in relatively long periods of anticipation. (ii) The anticipated representations may be insufficient to reflect the desired semantics of the target activity, due to lack of contextual clues. To address these issues, we propose "HRO", a hybrid framework that integrates both the memory-augmented recurrent and one-shot representation forecasting strategies.

1754, TITLE: It's All in the Teacher: Zero-Shot Quantization Brought Closer to the Teacher
http://openaccess.thecvf.com//content/CVPR2022/html/Choi_Its_All_in_the_Teacher_Zero-Shot_Quantization_Brought_Closer_to_CVPR_2022_paper.html
AUTHORS: Kanghyun Choi, Hye Yoon Lee, Deokki Hong, Joonsang Yu, Noseong Park, Youngsok Kim, Jinho Lee
HIGHLIGHT: Based on the observations, we propose AIT, a simple yet powerful technique for zero-shot quantization, which addresses the aforementioned two problems in the following way: AIT i) uses a KL distance loss only without a cross-entropy loss, and ii) manipulates gradients to guarantee that a certain portion of weights are properly updated after crossing the rounding thresholds.

1755, TITLE: Improving Segmentation of the Inferior Alveolar Nerve Through Deep Label Propagation
http://openaccess.thecvf.com//content/CVPR2022/html/Cipriano_Improving_Segmentation_of_the_Inferior_Alveolar_Nerve_Through_Deep_Label_CVPR_2022_paper.html
AUTHORS: Marco Cipriano, Stefano Allegretti, Federico Bolelli, Federico Pollastri, Costantino Grana
HIGHLIGHT: This paper exploits and publicly releases a new 3D densely annotated dataset, through which we are able to train a deep label propagation model which obtains better results than those available in literature.

1756, TITLE: A Text Attention Network for Spatial Deformation Robust Scene Text Image Super-Resolution
http://openaccess.thecvf.com//content/CVPR2022/html/Ma_A_Text_Attention_Network_for_Spatial_Deformation_Robust_Scene_Text_CVPR_2022_paper.html
AUTHORS: Jianqi Ma, Zhetong Liang, Lei Zhang
HIGHLIGHT: This is because the current CNN-based methods adopt locality-based operations, which are not effective to deal with the variation caused by deformations. In this paper, we propose a CNN based Text ATTention network (TATT) to address this problem.

1757, TITLE: Multi-Modal Dynamic Graph Transformer for Visual Grounding
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Multi-Modal_Dynamic_Graph_Transformer_for_Visual_Grounding_CVPR_2022_paper.html
AUTHORS: Sijia Chen, Baochun Li
HIGHLIGHT: Their performance depends on the density and quality of the candidate regions and is capped by the inability to optimize the located regions continuously. To address these issues, we propose to remodel VG into a progressively optimized visual semantic alignment process.

1758, TITLE: OSOP: A Multi-Stage One Shot Object Pose Estimation Framework
http://openaccess.thecvf.com//content/CVPR2022/html/Shugurov_OSOP_A_Multi-Stage_One_Shot_Object_Pose_Estimation_Framework_CVPR_2022_paper.html
AUTHORS: Ivan Shugurov, Fu Li, Benjamin Busam, Slobodan Ilic
HIGHLIGHT: We present a novel one-shot method for object detection and 6 DoF pose estimation, that does not require training on target objects.

1759, TITLE: Generative Cooperative Learning for Unsupervised Video Anomaly Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Zaheer_Generative_Cooperative_Learning_for_Unsupervised_Video_Anomaly_Detection_CVPR_2022_paper.html
AUTHORS: M. Zaigham Zaheer, Arif Mahmood, M. Haris Khan, Mattia Segu, Fisher Yu, Seung-Ik Lee
HIGHLIGHT: This problem is challenging yet rewarding as it can completely eradicate the costs of obtaining laborious annotations and enable such systems to be deployed without human intervention. To this end, we propose a novel unsupervised Generative Cooperative Learning (GCL) approach for video anomaly detection that exploits the low frequency of anomalies towards building a cross-supervision between a generator and a discriminator.

1760, TITLE: Rethinking Semantic Segmentation: A Prototype View
http://openaccess.thecvf.com//content/CVPR2022/html/Zhou_Rethinking_Semantic_Segmentation_A_Prototype_View_CVPR_2022_paper.html
AUTHORS: Tianfei Zhou, Wenguan Wang, Ender Konukoglu, Luc Van Gool
HIGHLIGHT: Prevalent semantic segmentation solutions, despite their different network designs (FCN based or attention based) and mask decoding strategies (parametric softmax based or pixel-query based), can be placed in one category, by considering the softmax weights or query vectors as learnable class prototypes. In light of this prototype view, this study uncovers several limitations of such parametric segmentation regime, and proposes a nonparametric alternative based on non-learnable prototypes.

1761, TITLE: Geometric Transformer for Fast and Robust Point Cloud Registration
http://openaccess.thecvf.com//content/CVPR2022/html/Qin_Geometric_Transformer_for_Fast_and_Robust_Point_Cloud_Registration_CVPR_2022_paper.html
AUTHORS: Zheng Qin, Hao Yu, Changjian Wang, Yulan Guo, Yuxing Peng, Kai Xu

HIGHLIGHT: We propose Geometric Transformer to learn geometric feature for robust superpoint matching.

1762, TITLE: Cross-Model Pseudo-Labeling for Semi-Supervised Action Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_Cross-Model_Pseudo-Labeling_for_Semi-Supervised_Action_Recognition_CVPR_2022_paper.html
AUTHORS: Yinghao Xu, Fangyun Wei, Xiao Sun, Ceyuan Yang, Yujun Shen, Bo Dai, Bolei Zhou, Stephen Lin
HIGHLIGHT: Typically in recent work, the pseudo-labels are obtained by training a model on the labeled data, and then using confident predictions from the model to teach itself.In this work, we propose a more effective pseudo-labeling scheme, called Cross-Model Pseudo-Labeling (CMPL).

1763, TITLE: UMT: Unified Multi-Modal Transformers for Joint Video Moment Retrieval and Highlight Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_UMT_Unified_Multi-Modal_Transformers_for_Joint_Video_Moment_Retrieval_and_CVPR_2022_paper.html
AUTHORS: Ye Liu, Siyuan Li, Yang Wu, Chang-Wen Chen, Ying Shan, Xiaohu Qie
HIGHLIGHT: Nevertheless, jointly conducting moment retrieval and highlight detection is an emerging research topic, even though its component problems and some related tasks have already been studied for a while. In this paper, we present the first unified framework, named Unified Multi-modal Transformers (UMT), capable of realizing such joint optimization while can also be easily degenerated for solving individual problems.

1764, TITLE: Dual-Shutter Optical Vibration Sensing
http://openaccess.thecvf.com//content/CVPR2022/html/Sheinin_Dual-Shutter_Optical_Vibration_Sensing_CVPR_2022_paper.html
AUTHORS: Mark Sheinin, Dorian Chan, Matthew O'Toole, Srinivasa G. Narasimhan
HIGHLIGHT: In this paper, we propose a novel method for sensing vibrations at high speeds (up to 63kHz), for multiple scene sources at once, using sensors rated for only 130Hz operation.

1765, TITLE: Demystifying the Neural Tangent Kernel From a Practical Perspective: Can It Be Trusted for Neural Architecture Search Without Training?
http://openaccess.thecvf.com//content/CVPR2022/html/Mok_Demystifying_the_Neural_Tangent_Kernel_From_a_Practical_Perspective_Can_CVPR_2022_paper.html
AUTHORS: Jisoo Mok, Byunggook Na, Ji-Hoon Kim, Dongyoon Han, Sungroh Yoon
HIGHLIGHT: In this work, we revisit several at-initialization metrics that can be derived from the NTK and reveal their key shortcomings.

1766, TITLE: Learning To Find Good Models in RANSAC
http://openaccess.thecvf.com//content/CVPR2022/html/Barath_Learning_To_Find_Good_Models_in_RANSAC_CVPR_2022_paper.html
AUTHORS: Daniel Barath, Luca Cavalli, Marc Pollefeys
HIGHLIGHT: We propose the Model Quality Network, MQ-Net in short, for predicting the quality, e.g. the pose error of essential matrices, of models generated inside RANSAC.

1767, TITLE: Interactiveness Field in Human-Object Interactions
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Interactiveness_Field_in_Human-Object_Interactions_CVPR_2022_paper.html
AUTHORS: Xinpeng Liu, Yong-Lu Li, Xiaoqian Wu, Yu-Wing Tai, Cewu Lu, Chi-Keung Tang
HIGHLIGHT: In this work, we introduce a previously overlooked interactiveness bimodal prior: given an object in an image, after pairing it with the humans, the generated pairs are either mostly non-interactive, or mostly interactive, with the former more frequent than the latter.

1768, TITLE: BodyGAN: General-Purpose Controllable Neural Human Body Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_BodyGAN_General-Purpose_Controllable_Neural_Human_Body_Generation_CVPR_2022_paper.html
AUTHORS: Chaojie Yang, Hanhui Li, Shengjie Wu, Shengkai Zhang, Haonan Yan, Nianhong Jiao, Jie Tang, Runnan Zhou, Xiaodan Liang, Tianxiang Zheng
HIGHLIGHT: In addition, such a unimodal strategy is prone to causing severe artifacts in the generated images like color distortions and unrealistic textures. To tackle these issues, this paper proposes a multi-factor conditioned method dubbed BodyGAN.

1769, TITLE: Image Disentanglement Autoencoder for Steganography Without Embedding
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Image_Disentanglement_Autoencoder_for_Steganography_Without_Embedding_CVPR_2022_paper.html
AUTHORS: Xiyao Liu, Ziping Ma, Junxing Ma, Jian Zhang, Gerald Schaefer, Hui Fang
HIGHLIGHT: In this paper, we propose Image DisEntanglement Autoencoder for Steganography (IDEAS) as a novel steganography without embedding (SWE) technique.

1770, TITLE: Self-Supervised Dense Consistency Regularization for Image-to-Image Translation
http://openaccess.thecvf.com//content/CVPR2022/html/Ko_Self-Supervised_Dense_Consistency_Regularization_for_Image-to-Image_Translation_CVPR_2022_paper.html
AUTHORS: Minsu Ko, Eunju Cha, Sungjoo Suh, Huijin Lee, Jae-Joon Han, Jinwoo Shin, Bohyung Han
HIGHLIGHT: In this paper, we present a simple but effective regularization technique for improving GAN-based image-to-image translation.

188

1771, TITLE: The Devil Is in the Details: Window-Based Attention for Image Compression
http://openaccess.thecvf.com//content/CVPR2022/html/Zou_The_Devil_Is_in_the_Details_Window-Based_Attention_for_Image_CVPR_2022_paper.html
AUTHORS: Renjie Zou, Chunfeng Song, Zhaoxiang Zhang
HIGHLIGHT: In this paper, we first extensively study the effects of multiple kinds of attention mechanisms for local features learning, then introduce a more straightforward yet effective window-based local attention block.

1772, TITLE: Category-Aware Transformer Network for Better Human-Object Interaction Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Dong_Category-Aware_Transformer_Network_for_Better_Human-Object_Interaction_Detection_CVPR_2022_paper.html
AUTHORS: Leizhen Dong, Zhimin Li, Kunlun Xu, Zhijun Zhang, Luxin Yan, Sheng Zhong, Xu Zou
HIGHLIGHT: In this paper, we try to study the issue of promoting transformerbased HOI detectors by initializing the Object Query with category-aware semantic information.

1773, TITLE: Deep Depth From Focus With Differential Focus Volume
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Deep_Depth_From_Focus_With_Differential_Focus_Volume_CVPR_2022_paper.html
AUTHORS: Fengting Yang, Xiaolei Huang, Zihan Zhou
HIGHLIGHT: In this work, we propose a convolutional neural network (CNN) to find the best-focused pixels in a focal stack and infer depth from the focus estimation.

1774, TITLE: DiLiGenT102: A Photometric Stereo Benchmark Dataset With Controlled Shape and Material Variation
http://openaccess.thecvf.com//content/CVPR2022/html/Ren_DiLiGenT102_A_Photometric_Stereo_Benchmark_Dataset_With_Controlled_Shape_and_CVPR_2022_paper.html
AUTHORS: Jieji Ren, Feishi Wang, Jiahao Zhang, Qian Zheng, Mingjun Ren, Boxin Shi
HIGHLIGHT: This paper presents a new real-world photometric stereo dataset with "ground truth" normal maps, which is 10 times larger than the widely adopted one.

1775, TITLE: Robust Fine-Tuning of Zero-Shot Models
http://openaccess.thecvf.com//content/CVPR2022/html/Wortsman_Robust_Fine-Tuning_of_Zero-Shot_Models_CVPR_2022_paper.html
AUTHORS: Mitchell Wortsman, Gabriel Ilharco, Jong Wook Kim, Mike Li, Simon Kornblith, Rebecca Roelofs, Raphael Gontijo Lopes, Hannaneh Hajishirzi, Ali Farhadi, Hongseok Namkoong, Ludwig Schmidt
HIGHLIGHT: Although existing fine-tuning methods substantially improve accuracy on a given target distribution, they often reduce robustness to distribution shifts. We address this tension by introducing a simple and effective method for improving robustness while fine-tuning: ensembling the weights of the zero-shot and fine-tuned models (WiSE-FT).

1776, TITLE: Towards Data-Free Model Stealing in a Hard Label Setting
http://openaccess.thecvf.com//content/CVPR2022/html/Sanyal_Towards_Data-Free_Model_Stealing_in_a_Hard_Label_Setting_CVPR_2022_paper.html
AUTHORS: Sunandini Sanyal, Sravanti Addepalli, R. Venkatesh Babu
HIGHLIGHT: In this work, we show that it is indeed possible to steal Machine Learning models by accessing only top-1 predictions (Hard Label setting) as well, without access to model gradients (Black-Box setting) or even the training dataset (Data-Free setting) within a low query budget.

1777, TITLE: PolyWorld: Polygonal Building Extraction With Graph Neural Networks in Satellite Images
http://openaccess.thecvf.com//content/CVPR2022/html/Zorzi_PolyWorld_Polygonal_Building_Extraction_With_Graph_Neural_Networks_in_Satellite_CVPR_2022_paper.html
AUTHORS: Stefano Zorzi, Shabab Bazrafkan, Stefan Habenschuss, Friedrich Fraundorfer
HIGHLIGHT: This paper introduces PolyWorld, a neural network that directly extracts building vertices from an image and connects them correctly to create precise polygons.

1778, TITLE: GAT-CADNet: Graph Attention Network for Panoptic Symbol Spotting in CAD Drawings
http://openaccess.thecvf.com//content/CVPR2022/html/Zheng_GAT-CADNet_Graph_Attention_Network_for_Panoptic_Symbol_Spotting_in_CAD_CVPR_2022_paper.html
AUTHORS: Zhaohua Zheng, Jianfang Li, Lingjie Zhu, Honghua Li, Frank Petzold, Ping Tan
HIGHLIGHT: By treating each CAD drawing as a graph, we propose a novel graph attention network GAT-CADNet to solve the panoptic symbol spotting problem: vertex features derived from the GAT branch are mapped to semantic labels, while their attention scores are cascaded and mapped to instance prediction.

1779, TITLE: Multi-Granularity Alignment Domain Adaptation for Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Zhou_Multi-Granularity_Alignment_Domain_Adaptation_for_Object_Detection_CVPR_2022_paper.html
AUTHORS: Wenzhang Zhou, Dawei Du, Libo Zhang, Tiejian Luo, Yanjun Wu
HIGHLIGHT: In this paper, we propose a unified multi-granularity alignment based object detection framework towards domain-invariant feature learning.

1780, TITLE: LARGE: Latent-Based Regression Through GAN Semantics
http://openaccess.thecvf.com//content/CVPR2022/html/Nitzan_LARGE_Latent-Based_Regression_Through_GAN_Semantics_CVPR_2022_paper.html

AUTHORS: Yotam Nitzan, Rinon Gal, Ofir Brenner, Daniel Cohen-Or
HIGHLIGHT: We propose a novel method for solving regression tasks using few-shot or weak supervision.

1781, TITLE: Are Multimodal Transformers Robust to Missing Modality?
http://openaccess.thecvf.com//content/CVPR2022/html/Ma_Are_Multimodal_Transformers_Robust_to_Missing_Modality_CVPR_2022_paper.html
AUTHORS: Mengmeng Ma, Jian Ren, Long Zhao, Davide Testuggine, Xi Peng
HIGHLIGHT: What surprised us is that the optimal fusion strategy is dataset dependent even for the same Transformer model; there does not exist a universal strategy that works in general cases. Based on these findings, we propose a principle method to improve the robustness of Transformer models by automatically searching for an optimal fusion strategy regarding input data.

1782, TITLE: Degradation-Agnostic Correspondence From Resolution-Asymmetric Stereo
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Degradation-Agnostic_Correspondence_From_Resolution-Asymmetric_Stereo_CVPR_2022_paper.html
AUTHORS: Xihao Chen, Zhiwei Xiong, Zhen Cheng, Jiayong Peng, Yueyi Zhang, Zheng-Jun Zha
HIGHLIGHT: In this paper, we study the problem of stereo matching from a pair of images with different resolutions, e.g., those acquired with a tele-wide camera system.

1783, TITLE: Fisher Information Guidance for Learned Time-of-Flight Imaging
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Fisher_Information_Guidance_for_Learned_Time-of-Flight_Imaging_CVPR_2022_paper.html
AUTHORS: Jiaqu Li, Tao Yue, Sijie Zhao, Xuemei Hu
HIGHLIGHT: In this paper, we propose a Fisher-information guided framework to jointly optimize the coding functions (light modulation and sensor demodulation functions) and the reconstruction network of iToF imaging, with the supervision of the proposed discriminative fisher loss.

1784, TITLE: VRDFormer: End-to-End Video Visual Relation Detection With Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Zheng_VRDFormer_End-to-End_Video_Visual_Relation_Detection_With_Transformers_CVPR_2022_paper.html
AUTHORS: Sipeng Zheng, Shizhe Chen, Qin Jin
HIGHLIGHT: Most previous works adopt a multi-stage framework for video visual relation detection (VidVRD), which cannot capture long-term spatiotemporal contexts in different stages and also suffers from inefficiency. In this paper, we propose a transformerbased framework called VRDFormer to unify these decoupling stages.

1785, TITLE: Robust Federated Learning With Noisy and Heterogeneous Clients
http://openaccess.thecvf.com//content/CVPR2022/html/Fang_Robust_Federated_Learning_With_Noisy_and_Heterogeneous_Clients_CVPR_2022_paper.html
AUTHORS: Xiuwen Fang, Mang Ye
HIGHLIGHT: We present a novel solution RHFL (Robust Heterogeneous Federated Learning), which simultaneously handles the label noise and performs federated learning in a single framework.

1786, TITLE: Enabling Equivariance for Arbitrary Lie Groups
http://openaccess.thecvf.com//content/CVPR2022/html/MacDonald_Enabling_Equivariance_for_Arbitrary_Lie_Groups_CVPR_2022_paper.html
AUTHORS: Lachlan E. MacDonald, Sameera Ramasinghe, Simon Lucey
HIGHLIGHT: We introduce a rigourous mathematical framework to permit invariance to any Lie group of warps, exclusively using convolutions (over Lie groups), without the need for capsules.

1787, TITLE: Unbiased Teacher v2: Semi-Supervised Object Detection for Anchor-Free and Anchor-Based Detectors
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Unbiased_Teacher_v2_Semi-Supervised_Object_Detection_for_Anchor-Free_and_Anchor-Based_CVPR_2022_paper.html
AUTHORS: Yen-Cheng Liu, Chih-Yao Ma, Zsolt Kira
HIGHLIGHT: In this paper, we present Unbiased Teacher v2, which shows the generalization of SS-OD method to anchor-free detectors and also introduces Listen2Student mechanism for the unsupervised regression loss.

1788, TITLE: GPU-Based Homotopy Continuation for Minimal Problems in Computer Vision
http://openaccess.thecvf.com//content/CVPR2022/html/Chien_GPU-Based_Homotopy_Continuation_for_Minimal_Problems_in_Computer_Vision_CVPR_2022_paper.html
AUTHORS: Chiang-Heng Chien, Hongyi Fan, Ahmad Abdelfattah, Elias Tsigaridas, Stanimire Tomov, Benjamin Kimia
HIGHLIGHT: In this paper we show that HC can be parallelized on a GPU, showing significant speedups up to 56 times on polynomial benchmarks.

1789, TITLE: Learning Pixel-Level Distinctions for Video Highlight Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Wei_Learning_Pixel-Level_Distinctions_for_Video_Highlight_Detection_CVPR_2022_paper.html
AUTHORS: Fanyue Wei, Biao Wang, Tiezheng Ge, Yuning Jiang, Wen Li, Lixin Duan
HIGHLIGHT: To this end, we propose to learn pixel-level distinctions to improve the video highlight detection.

1790, TITLE: Noise Distribution Adaptive Self-Supervised Image Denoising Using Tweedie Distribution and Score Matching

190

http://openaccess.thecvf.com//content/CVPR2022/html/Kim_Noise_Distribution_Adaptive_Self-Supervised_Image_Denoising_Using_Tweedie_Distribution_and_CVPR_2022_paper.html
AUTHORS: Kwanyoung Kim, Taesung Kwon, Jong Chul Ye
HIGHLIGHT: Here, we reveal that Tweedie distributions also play key roles in modern deep learning era, leading to a distribution independent self-supervised image denoising formula without clean reference images.

1791, TITLE: Lite Pose: Efficient Architecture Design for 2D Human Pose Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Lite_Pose_Efficient_Architecture_Design_for_2D_Human_Pose_Estimation_CVPR_2022_paper.html
AUTHORS: Yihan Wang, Muyang Li, Han Cai, Wei-Ming Chen, Song Han
HIGHLIGHT: In this paper, we study efficient architecture design for real-time multi-person pose estimation on edge.

1792, TITLE: Boosting Black-Box Attack With Partially Transferred Conditional Adversarial Distribution
http://openaccess.thecvf.com//content/CVPR2022/html/Feng_Boosting_Black-Box_Attack_With_Partially_Transferred_Conditional_Adversarial_Distribution_CVPR_2022_paper.html
AUTHORS: Yan Feng, Baoyuan Wu, Yanbo Fan, Li Liu, Zhifeng Li, Shu-Tao Xia
HIGHLIGHT: However, due to the possible differences on model architectures and training datasets between surrogate and target models, dubbed "surrogate biases", the contribution of adversarial transferability to improving the attack performance may be weakened. To tackle this issue, we innovatively propose a black-box attack method by developing a novel mechanism of adversarial transferability, which is robust to the surrogate biases.

1793, TITLE: CLIPstyler: Image Style Transfer With a Single Text Condition
http://openaccess.thecvf.com//content/CVPR2022/html/Kwon_CLIPstyler_Image_Style_Transfer_With_a_Single_Text_Condition_CVPR_2022_paper.html
AUTHORS: Gihyun Kwon, Jong Chul Ye
HIGHLIGHT: However, in many practical situations, users may not have reference style images but still be interested in transferring styles by just imagining them. In order to deal with such applications, we propose a new framework that enables a style transfer 'without' a style image, but only with a text description of the desired style.

1794, TITLE: Ray Priors Through Reprojection: Improving Neural Radiance Fields for Novel View Extrapolation
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Ray_Priors_Through_Reprojection_Improving_Neural_Radiance_Fields_for_Novel_CVPR_2022_paper.html
AUTHORS: Jian Zhang, Yuanqing Zhang, Huan Fu, Xiaowei Zhou, Bowen Cai, Jinchi Huang, Rongfei Jia, Binqiang Zhao, Xing Tang
HIGHLIGHT: In this paper, instead of exploiting few-shot image synthesis, we study the novel view extrapolation setting that (1) the training images can well describe an object, and (2) there is a notable discrepancy between the training and test viewpoints' distributions.

1795, TITLE: Spatio-Temporal Relation Modeling for Few-Shot Action Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Thatipelli_Spatio-Temporal_Relation_Modeling_for_Few-Shot_Action_Recognition_CVPR_2022_paper.html
AUTHORS: Anirudh Thatipelli, Sanath Narayan, Salman Khan, Rao Muhammad Anwer, Fahad Shahbaz Khan, Bernard Ghanem
HIGHLIGHT: We propose a novel few-shot action recognition framework, STRM, which enhances class-specific feature discriminability while simultaneously learning higher-order temporal representations.

1796, TITLE: Pop-Out Motion: 3D-Aware Image Deformation via Learning the Shape Laplacian
http://openaccess.thecvf.com//content/CVPR2022/html/Lee_Pop-Out_Motion_3D-Aware_Image_Deformation_via_Learning_the_Shape_Laplacian_CVPR_2022_paper.html
AUTHORS: Jihyun Lee, Minhyuk Sung, Hyunjin Kim, Tae-Kyun Kim
HIGHLIGHT: We propose a framework that can deform an object in a 2D image as it exists in 3D space.

1797, TITLE: Volumetric Bundle Adjustment for Online Photorealistic Scene Capture
http://openaccess.thecvf.com//content/CVPR2022/html/Clark_Volumetric_Bundle_Adjustment_for_Online_Photorealistic_Scene_Capture_CVPR_2022_paper.html
AUTHORS: Ronald Clark
HIGHLIGHT: In this paper, we propose a system that can reconstruct photorealistic models of complex scenes in an efficient manner.

1798, TITLE: Multi-Person Extreme Motion Prediction
http://openaccess.thecvf.com//content/CVPR2022/html/Guo_Multi-Person_Extreme_Motion_Prediction_CVPR_2022_paper.html
AUTHORS: Wen Guo, Xiaoyu Bie, Xavier Alameda-Pineda, Francesc Moreno-Noguer
HIGHLIGHT: While this problem has recently received increasing attention, it has mostly been tackled for single humans in isolation. In this paper, we explore this problem when dealing with humans performing collaborative tasks, we seek to predict the future motion of two interacted persons given two sequences of their past skeletons.

1799, TITLE: Masking Adversarial Damage: Finding Adversarial Saliency for Robust and Sparse Network
http://openaccess.thecvf.com//content/CVPR2022/html/Lee_Masking_Adversarial_Damage_Finding_Adversarial_Saliency_for_Robust_and_Sparse_CVPR_2022_paper.html
AUTHORS: Byung-Kwan Lee, Junho Kim, Yong Man Ro

HIGHLIGHT: To bridge adversarial robustness and model compression, we propose a novel adversarial pruning method, Masking Adversarial Damage (MAD) that employs second-order information of adversarial loss.

1800, TITLE: Channel Balancing for Accurate Quantization of Winograd Convolutions
http://openaccess.thecvf.com//content/CVPR2022/html/Chikin_Channel_Balancing_for_Accurate_Quantization_of_Winograd_Convolutions_CVPR_2022_paper.html
AUTHORS: Vladimir Chikin, Vladimir Kryzhanovskiy
HIGHLIGHT: We introduce a novel class of Winograd algorithms that balances the filter and input channels in the Winograd domain.

1801, TITLE: RegNeRF: Regularizing Neural Radiance Fields for View Synthesis From Sparse Inputs
http://openaccess.thecvf.com//content/CVPR2022/html/Niemeyer_RegNeRF_Regularizing_Neural_Radiance_Fields_for_View_Synthesis_From_Sparse_CVPR_2022_paper.html
AUTHORS: Michael Niemeyer, Jonathan T. Barron, Ben Mildenhall, Mehdi S. M. Sajjadi, Andreas Geiger, Noha Radwan
HIGHLIGHT: We observe that the majority of artifacts in sparse input scenarios are caused by errors in the estimated scene geometry, and by divergent behavior at the start of training. We address this by regularizing the geometry and appearance of patches rendered from unobserved viewpoints, and annealing the ray sampling space during training.

1802, TITLE: Structured Local Radiance Fields for Human Avatar Modeling
http://openaccess.thecvf.com//content/CVPR2022/html/Zheng_Structured_Local_Radiance_Fields_for_Human_Avatar_Modeling_CVPR_2022_paper.html
AUTHORS: Zerong Zheng, Han Huang, Tao Yu, Hongwen Zhang, Yandong Guo, Yebin Liu
HIGHLIGHT: To learn our representation from RGB data and facilitate pose generalization, we propose to learn the node translations and the detail variations in a conditional generative latent space.

1803, TITLE: Towards Noiseless Object Contours for Weakly Supervised Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Towards_Noiseless_Object_Contours_for_Weakly_Supervised_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Jing Li, Junsong Fan, Zhaoxiang Zhang
HIGHLIGHT: In this paper, we train a SANCE model which utilizes an auxiliary segmentation module to supplement high-level semantic information for contour training by backbone feature sharing and online label supervision.

1804, TITLE: Ranking-Based Siamese Visual Tracking
http://openaccess.thecvf.com//content/CVPR2022/html/Tang_Ranking-Based_Siamese_Visual_Tracking_CVPR_2022_paper.html
AUTHORS: Feng Tang, Qiang Ling
HIGHLIGHT: Moreover, such tracking paradigm takes only the classification confidence of proposals for the final prediction, which may yield the misalignment between classification and localization. To resolve these issues, this paper proposes a ranking-based optimization algorithm to explore the relationship among different proposals.

1805, TITLE: Learnable Lookup Table for Neural Network Quantization
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Learnable_Lookup_Table_for_Neural_Network_Quantization_CVPR_2022_paper.html
AUTHORS: Longguang Wang, Xiaoyu Dong, Yingqian Wang, Li Liu, Wei An, Yulan Guo
HIGHLIGHT: In this paper, we formulate the quantization process as a simple lookup operation and propose to learn lookup tables as quantizers.

1806, TITLE: SEEG: Semantic Energized Co-Speech Gesture Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Liang_SEEG_Semantic_Energized_Co-Speech_Gesture_Generation_CVPR_2022_paper.html
AUTHORS: Yuanzhi Liang, Qianyu Feng, Linchao Zhu, Li Hu, Pan Pan, Yi Yang
HIGHLIGHT: In this paper, we propose a novel method SEmantic Energized Generation (SEEG), for semantic-aware gesture generation.

1807, TITLE: AdaViT: Adaptive Vision Transformers for Efficient Image Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Meng_AdaViT_Adaptive_Vision_Transformers_for_Efficient_Image_Recognition_CVPR_2022_paper.html
AUTHORS: Lingchen Meng, Hengduo Li, Bor-Chun Chen, Shiyi Lan, Zuxuan Wu, Yu-Gang Jiang, Ser-Nam Lim
HIGHLIGHT: In this paper, we argue that due to the large variations among images, their need for modeling long-range dependencies between patches differ.

1808, TITLE: Compound Domain Generalization via Meta-Knowledge Encoding
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Compound_Domain_Generalization_via_Meta-Knowledge_Encoding_CVPR_2022_paper.html
AUTHORS: Chaoqi Chen, Jiongcheng Li, Xiaoguang Han, Xiaoqing Liu, Yizhou Yu
HIGHLIGHT: In this paper, we study a practical problem of compound DG, which relaxes the discrete domain assumption to the mixed source domains setting.

1809, TITLE: NAN: Noise-Aware NeRFs for Burst-Denoising
http://openaccess.thecvf.com//content/CVPR2022/html/Pearl_NAN_Noise-Aware_NeRFs_for_Burst-Denoising_CVPR_2022_paper.html

AUTHORS: Naama Pearl, Tali Treibitz, Simon Korman
HIGHLIGHT: We show that Neural Radiance Fields (NeRFs), originally suggested for physics-based novel-view rendering, can serve as a powerful framework for burst denoising.

1810, TITLE: Physical Inertial Poser (PIP): Physics-Aware Real-Time Human Motion Tracking From Sparse Inertial Sensors
http://openaccess.thecvf.com//content/CVPR2022/html/Yi_Physical_Inertial_Poser_PIP_Physics-Aware_Real-Time_Human_Motion_Tracking_From_CVPR_2022_paper.html
AUTHORS: Xinyu Yi, Yuxiao Zhou, Marc Habermann, Soshi Shimada, Vladislav Golyanik, Christian Theobalt, Feng Xu
HIGHLIGHT: In consequence, recent state-of-the-art methods can barely handle very long period motions, and unrealistic artifacts are common due to the unawareness of physical constraints. To this end, we present the first method which combines a neural kinematics estimator and a physics-aware motion optimizer to track body motions with only 6 inertial sensors.

1811, TITLE: b-DARTS: Beta-Decay Regularization for Differentiable Architecture Search
http://openaccess.thecvf.com//content/CVPR2022/html/Ye_b-DARTS_Beta-Decay_Regularization_for_Differentiable_Architecture_Search_CVPR_2022_paper.html
AUTHORS: Peng Ye, Baopu Li, Yikang Li, Tao Chen, Jiayuan Fan, Wanli Ouyang
HIGHLIGHT: However, they suffer from two main issues, the weak robustness to the performance collapse and the poor generalization ability of the searched architectures. To solve these two problems, a simple-but-efficient regularization method, termed as Beta-Decay, is proposed to regularize the DARTS-based NAS searching process.

1812, TITLE: Vector Quantized Diffusion Model for Text-to-Image Synthesis
http://openaccess.thecvf.com//content/CVPR2022/html/Gu_Vector_Quantized_Diffusion_Model_for_Text-to-Image_Synthesis_CVPR_2022_paper.html
AUTHORS: Shuyang Gu, Dong Chen, Jianmin Bao, Fang Wen, Bo Zhang, Dongdong Chen, Lu Yuan, Baining Guo
HIGHLIGHT: We present the vector quantized diffusion (VQ-Diffusion) model for text-to-image generation.

1813, TITLE: CMT: Convolutional Neural Networks Meet Vision Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Guo_CMT_Convolutional_Neural_Networks_Meet_Vision_Transformers_CVPR_2022_paper.html
AUTHORS: Jianyuan Guo, Kai Han, Han Wu, Yehui Tang, Xinghao Chen, Yunhe Wang, Chang Xu
HIGHLIGHT: However, there are still gaps in both performance and computational cost between transformers and existing convolutional neural networks (CNNs). In this paper, we aim to address this issue and develop a network that can outperform not only the canonical transformers, but also the high-performance convolutional models.

1814, TITLE: Hyperspherical Consistency Regularization
http://openaccess.thecvf.com//content/CVPR2022/html/Tan_Hyperspherical_Consistency_Regularization_CVPR_2022_paper.html
AUTHORS: Cheng Tan, Zhangyang Gao, Lirong Wu, Siyuan Li, Stan Z. Li
HIGHLIGHT: In this work, we systematically explore the relationship between self-supervised learning and supervised learning, and study how self-supervised learning helps robust data-efficient deep learning.

1815, TITLE: Unsupervised Image-to-Image Translation With Generative Prior
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Unsupervised_Image-to-Image_Translation_With_Generative_Prior_CVPR_2022_paper.html
AUTHORS: Shuai Yang, Liming Jiang, Ziwei Liu, Chen Change Loy
HIGHLIGHT: In this work, we present a novel framework, Generative Prior-guided UNsupervised Image-to-image Translation (GP-UNIT), to improve the overall quality and applicability of the translation algorithm.

1816, TITLE: KNN Local Attention for Image Restoration
http://openaccess.thecvf.com//content/CVPR2022/html/Lee_KNN_Local_Attention_for_Image_Restoration_CVPR_2022_paper.html
AUTHORS: Hunsang Lee, Hyesong Choi, Kwanghoon Sohn, Dongbo Min
HIGHLIGHT: However, by focusing only on adjacent positions, the local attention suffers from an insufficient receptive field for image restoration. In this paper, we propose a new attention mechanism for image restoration, called k-NN Image Transformer (KiT), that rectifies above mentioned limitations.

1817, TITLE: Face Relighting With Geometrically Consistent Shadows
http://openaccess.thecvf.com//content/CVPR2022/html/Hou_Face_Relighting_With_Geometrically_Consistent_Shadows_CVPR_2022_paper.html
AUTHORS: Andrew Hou, Michel Sarkis, Ning Bi, Yiying Tong, Xiaoming Liu
HIGHLIGHT: We propose a novel differentiable algorithm for synthesizing hard shadows based on ray tracing, which we incorporate into training our face relighting model.

1818, TITLE: Open-Set Text Recognition via Character-Context Decoupling
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Open-Set_Text_Recognition_via_Character-Context_Decoupling_CVPR_2022_paper.html
AUTHORS: Chang Liu, Chun Yang, Xu-Cheng Yin
HIGHLIGHT: Under open-set scenarios, the intractable bias in contextual information can be passed down to visual information, consequently impairing the classification performance. In this paper, a Character-Context Decoupling framework is proposed to alleviate this problem by separating contextual information and character-visual information.

1819, TITLE: Multi-Marginal Contrastive Learning for Multi-Label Subcellular Protein Localization

http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Multi-Marginal_Contrastive_Learning_for_Multi-Label_Subcellular_Protein_Localization_CVPR_2022_paper.html
AUTHORS: Ziyi Liu, Zengmao Wang, Bo Du
HIGHLIGHT: In this paper, we propose a deep protein subcellular localization method with multi-marginal contrastive learning to perceive the same PSLs in different tissue images and different PSLs within the same tissue image.

1820, TITLE: Probabilistic Warp Consistency for Weakly-Supervised Semantic Correspondences
http://openaccess.thecvf.com//content/CVPR2022/html/Truong_Probabilistic_Warp_Consistency_for_Weakly-Supervised_Semantic_Correspondences_CVPR_2022_paper.html
AUTHORS: Prune Truong, Martin Danelljan, Fisher Yu, Luc Van Gool
HIGHLIGHT: We propose Probabilistic Warp Consistency, a weakly-supervised learning objective for semantic matching.

1821, TITLE: Predict, Prevent, and Evaluate: Disentangled Text-Driven Image Manipulation Empowered by Pre-Trained Vision-Language Model
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_Predict_Prevent_and_Evaluate_Disentangled_Text-Driven_Image_Manipulation_Empowered_by_CVPR_2022_paper.html
AUTHORS: Zipeng Xu, Tianwei Lin, Hao Tang, Fu Li, Dongliang He, Nicu Sebe, Radu Timofte, Luc Van Gool, Errui Ding
HIGHLIGHT: We propose a novel framework, i.e., Predict, Prevent, and Evaluate (PPE), for disentangled text-driven image manipulation that requires little manual annotation while being applicable to a wide variety of manipulations.

1822, TITLE: Optimizing Elimination Templates by Greedy Parameter Search
http://openaccess.thecvf.com//content/CVPR2022/html/Martyushev_Optimizing_Elimination_Templates_by_Greedy_Parameter_Search_CVPR_2022_paper.html
AUTHORS: Evgeniy Martyushev, Jana Vráblíková, Tomas Pajdla
HIGHLIGHT: We propose a new method for constructing elimination templates for efficient polynomial system solving of minimal problems in structure from motion, image matching, and camera tracking.

1823, TITLE: TransMix: Attend To Mix for Vision Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_TransMix_Attend_To_Mix_for_Vision_Transformers_CVPR_2022_paper.html
AUTHORS: Jie-Neng Chen, Shuyang Sun, Ju He, Philip H.S. Torr, Alan Yuille, Song Bai
HIGHLIGHT: This may lead to a strange phenomenon that sometimes there is no valid object in the mixed image due to the random process in augmentation but there is still response in the label space. To bridge such gap between the input and label spaces, we propose TransMix, which mixes labels based on the attention maps of Vision Transformers.

1824, TITLE: HOP: History-and-Order Aware Pre-Training for Vision-and-Language Navigation
http://openaccess.thecvf.com//content/CVPR2022/html/Qiao_HOP_History-and-Order_Aware_Pre-Training_for_Vision-and-Language_Navigation_CVPR_2022_paper.html
AUTHORS: Yanyuan Qiao, Yuankai Qi, Yicong Hong, Zheng Yu, Peng Wang, Qi Wu
HIGHLIGHT: In this work, to promote the learning of spatio-temporal visual-textual correspondence as well as the agent's capability of decision making, we propose a novel history-and-order aware pre-training paradigm (HOP) with VLN-specific objectives that exploit the past observations and support future action prediction.

1825, TITLE: Inertia-Guided Flow Completion and Style Fusion for Video Inpainting
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Inertia-Guided_Flow_Completion_and_Style_Fusion_for_Video_Inpainting_CVPR_2022_paper.html
AUTHORS: Kaidong Zhang, Jingjing Fu, Dong Liu
HIGHLIGHT: Nevertheless, the existing flow-guided cross-frame warping methods fail to consider the lightening and sharpness variation across video frames, which leads to spatial incoherence after warping from other frames. To alleviate such problem, we propose the Adaptive Style Fusion Network (ASFN), which utilizes the style information extracted from the valid regions to guide the gradient refinement in the warped regions.

1826, TITLE: RU-Net: Regularized Unrolling Network for Scene Graph Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Lin_RU-Net_Regularized_Unrolling_Network_for_Scene_Graph_Generation_CVPR_2022_paper.html
AUTHORS: Xin Lin, Changxing Ding, Jing Zhang, Yibing Zhan, Dacheng Tao
HIGHLIGHT: Existing SGG methods usually suffer from several issues, including 1) ambiguous object representations, as graph neural network-based message passing (GMP) modules are typically sensitive to spurious inter-node correlations, and 2) low diversity in relationship predictions due to severe class imbalance and a large number of missing annotations. To address both problems, in this paper, we propose a regularized unrolling network (RU-Net).

1827, TITLE: Long-Tailed Visual Recognition via Gaussian Clouded Logit Adjustment
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Long-Tailed_Visual_Recognition_via_Gaussian_Clouded_Logit_Adjustment_CVPR_2022_paper.html
AUTHORS: Mengke Li, Yiu-ming Cheung, Yang Lu
HIGHLIGHT: It is unfavorable for training on balanced data, but can be utilized to adjust the validity of the samples in long-tailed data, thereby solving the distorted embedding space of long-tailed problems. To this end, this paper proposes the Gaussian clouded logit adjustment by Gaussian perturbation of different class logits with varied amplitude.

1828, TITLE: Image Animation With Perturbed Masks

http://openaccess.thecvf.com//content/CVPR2022/html/Shalev_Image_Animation_With_Perturbed_Masks_CVPR_2022_paper.html
AUTHORS:　　　　　Yoav Shalev, Lior Wolf
HIGHLIGHT:　　　　We present a novel approach for image-animation of a source image by a driving video, both depicting the same type of object.

1829, TITLE:　　　　Exploring the Equivalence of Siamese Self-Supervised Learning via a Unified Gradient Framework
http://openaccess.thecvf.com//content/CVPR2022/html/Tao_Exploring_the_Equivalence_of_Siamese_Self-Supervised_Learning_via_a_Unified_CVPR_2022_paper.html
AUTHORS:　　　　　Chenxin Tao, Honghui Wang, Xizhou Zhu, Jiahua Dong, Shiji Song, Gao Huang, Jifeng Dai
HIGHLIGHT:　　　　The final accuracy numbers also vary, where different networks and tricks are utilized in different works. In this work, we demonstrate that these methods can be unified into the same form.

1830, TITLE:　　　　Point Density-Aware Voxels for LiDAR 3D Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Hu_Point_Density-Aware_Voxels_for_LiDAR_3D_Object_Detection_CVPR_2022_paper.html
AUTHORS:　　　　　Jordan S. K. Hu, Tianshu Kuai, Steven L. Waslander
HIGHLIGHT:　　　　Our proposed solution, Point Density-Aware Voxel network (PDV), is an end-to-end two stage LiDAR 3D object detection architecture that is designed to account for these point density variations.

1831, TITLE:　　　　Integrating Language Guidance Into Vision-Based Deep Metric Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Roth_Integrating_Language_Guidance_Into_Vision-Based_Deep_Metric_Learning_CVPR_2022_paper.html
AUTHORS:　　　　　Karsten Roth, Oriol Vinyals, Zeynep Akata
HIGHLIGHT:　　　　This causes learned embedding spaces to encode incomplete semantic context and misrepresent the semantic relation between classes, impacting the generalizability of the learned metric space. To tackle this issue, we propose a language guidance objective for visual similarity learning.

1832, TITLE:　　　　PartGlot: Learning Shape Part Segmentation From Language Reference Games
http://openaccess.thecvf.com//content/CVPR2022/html/Koo_PartGlot_Learning_Shape_Part_Segmentation_From_Language_Reference_Games_CVPR_2022_paper.html
AUTHORS:　　　　　Juil Koo, Ian Huang, Panos Achlioptas, Leonidas J. Guibas, Minhyuk Sung
HIGHLIGHT:　　　　We introduce PartGlot, a neural framework and associated architectures for learning semantic part segmentation of 3D shape geometry, based solely on part referential language.

1833, TITLE:　　　　Domain Generalization via Shuffled Style Assembly for Face Anti-Spoofing
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Domain_Generalization_via_Shuffled_Style_Assembly_for_Face_Anti-Spoofing_CVPR_2022_paper.html
AUTHORS:　　　　　Zhuo Wang, Zezheng Wang, Zitong Yu, Weihong Deng, Jiahong Li, Tingting Gao, Zhongyuan Wang
HIGHLIGHT:　　　　In this work, we separate the complete representation into content and style ones.

1834, TITLE:　　　　A Simple Episodic Linear Probe Improves Visual Recognition in the Wild
http://openaccess.thecvf.com//content/CVPR2022/html/Liang_A_Simple_Episodic_Linear_Probe_Improves_Visual_Recognition_in_the_CVPR_2022_paper.html
AUTHORS:　　　　　Yuanzhi Liang, Linchao Zhu, Xiaohan Wang, Yi Yang
HIGHLIGHT:　　　　In this paper, we propose an episodic linear probing (ELP) classifier to reflect the generalization of visual representations in an online manner.

1835, TITLE:　　　　Matching Feature Sets for Few-Shot Image Classification
http://openaccess.thecvf.com//content/CVPR2022/html/Afrasiyabi_Matching_Feature_Sets_for_Few-Shot_Image_Classification_CVPR_2022_paper.html
AUTHORS:　　　　　Arman Afrasiyabi, Hugo Larochelle, Jean-François Lalonde, Christian Gagné
HIGHLIGHT:　　　　Few-shot classification methods also mostly follow this trend. In this work, we depart from this established direction and instead propose to extract sets of feature vectors for each image.

1836, TITLE:　　　　DIVeR: Real-Time and Accurate Neural Radiance Fields With Deterministic Integration for Volume Rendering
http://openaccess.thecvf.com//content/CVPR2022/html/Wu_DIVeR_Real-Time_and_Accurate_Neural_Radiance_Fields_With_Deterministic_Integration_CVPR_2022_paper.html
AUTHORS:　　　　　Liwen Wu, Jae Yong Lee, Anand Bhattad, Yu-Xiong Wang, David Forsyth
HIGHLIGHT:　　　　DIVeR builds on the key ideas of NeRF and its variants -- density models and volume rendering -- to learn 3D object models that can be rendered realistically from small numbers of images.

1837, TITLE:　　　　Enhancing Classifier Conservativeness and Robustness by Polynomiality
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Enhancing_Classifier_Conservativeness_and_Robustness_by_Polynomiality_CVPR_2022_paper.html
AUTHORS:　　　　　Ziqi Wang, Marco Loog
HIGHLIGHT:　　　　We illustrate the detrimental effect, such as overconfident decisions, that exponential behavior can have in methods like classical LDA and logistic regression. We then show how polynomiality can remedy the situation.

1838, TITLE:　　　　Deep Spectral Methods: A Surprisingly Strong Baseline for Unsupervised Semantic Segmentation and Localization

http://openaccess.thecvf.com//content/CVPR2022/html/Melas-
Kyriazi_Deep_Spectral_Methods_A_Surprisingly_Strong_Baseline_for_Unsupervised_Semantic_CVPR_2022_paper.html
AUTHORS: Luke Melas-Kyriazi, Christian Rupprecht, Iro Laina, Andrea Vedaldi
HIGHLIGHT: Differently from existing methods, which are purely based on deep learning, we take inspiration from traditional spectral segmentation methods by reframing image decomposition as a graph partitioning problem.

1839, TITLE: OcclusionFusion: Occlusion-Aware Motion Estimation for Real-Time Dynamic 3D Reconstruction
http://openaccess.thecvf.com//content/CVPR2022/html/Lin_OcclusionFusion_Occlusion-Aware_Motion_Estimation_for_Real-Time_Dynamic_3D_Reconstruction_CVPR_2022_paper.html
AUTHORS: Wenbin Lin, Chengwei Zheng, Jun-Hai Yong, Feng Xu
HIGHLIGHT: This problem is even more severe for single-view-based systems due to strong occlusions. Based on these observations, we propose OcclusionFusion, a novel method to calculate occlusion-aware 3D motion to guide the reconstruction.

1840, TITLE: ContIG: Self-Supervised Multimodal Contrastive Learning for Medical Imaging With Genetics
http://openaccess.thecvf.com//content/CVPR2022/html/Taleb_ContIG_Self-Supervised_Multimodal_Contrastive_Learning_for_Medical_Imaging_With_Genetics_CVPR_2022_paper.html
AUTHORS: Aiham Taleb, Matthias Kirchler, Remo Monti, Christoph Lippert
HIGHLIGHT: In this work, we propose ContIG, a self-supervised method that can learn from large datasets of unlabeled medical images and genetic data.

1841, TITLE: Revisiting Domain Generalized Stereo Matching Networks From a Feature Consistency Perspective
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Revisiting_Domain_Generalized_Stereo_Matching_Networks_From_a_Feature_Consistency_CVPR_2022_paper.html
AUTHORS: Jiawei Zhang, Xiang Wang, Xiao Bai, Chen Wang, Lei Huang, Yimin Chen, Lin Gu, Jun Zhou, Tatsuya Harada, Edwin R. Hancock
HIGHLIGHT: We argue that maintaining feature consistency between matching pixels is a vital factor for promoting the generalization capability of stereo matching networks, which has not been adequately considered. Here we address this issue by proposing a simple pixel-wise contrastive learning across the viewpoints.

1842, TITLE: MonoScene: Monocular 3D Semantic Scene Completion
http://openaccess.thecvf.com//content/CVPR2022/html/Cao_MonoScene_Monocular_3D_Semantic_Scene_Completion_CVPR_2022_paper.html
AUTHORS: Anh-Quan Cao, Raoul de Charette
HIGHLIGHT: Along with architectural contributions, we introduce novel global scene and local frustums losses.

1843, TITLE: TubeFormer-DeepLab: Video Mask Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Kim_TubeFormer-DeepLab_Video_Mask_Transformer_CVPR_2022_paper.html
AUTHORS: Dahun Kim, Jun Xie, Huiyu Wang, Siyuan Qiao, Qihang Yu, Hong-Seok Kim, Hartwig Adam, In So Kweon, Liang-Chieh Chen
HIGHLIGHT: We present TubeFormer-DeepLab, the first attempt to tackle multiple core video segmentation tasks in a unified manner.

1844, TITLE: XMP-Font: Self-Supervised Cross-Modality Pre-Training for Few-Shot Font Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_XMP-Font_Self-Supervised_Cross-Modality_Pre-Training_for_Few-Shot_Font_Generation_CVPR_2022_paper.html
AUTHORS: Wei Liu, Fangyue Liu, Fei Ding, Qian He, Zili Yi
HIGHLIGHT: However, these few-shot font generation methods either fail to capture content-independent style representations, or employ localized component-wise style representations, which is insufficient to model many Chinese font styles that involve hyper-component features such as inter-component spacing and "connected-stroke". To resolve these drawbacks and make the style representations more reliable, we propose a self-supervised cross-modality pre-training strategy and a cross-modality transformer-based encoder that is conditioned jointly on the glyph image and the corresponding stroke labels.

1845, TITLE: Disentangling Visual and Written Concepts in CLIP
http://openaccess.thecvf.com//content/CVPR2022/html/Materzynska_Disentangling_Visual_and_Written_Concepts_in_CLIP_CVPR_2022_paper.html
AUTHORS: Joanna MaterzyÅ„ska, Antonio Torralba, David Bau
HIGHLIGHT: The CLIP network measures the similarity between natural text and images; in this work, we investigate the entanglement of the representation of word images and natural images in its image encoder.

1846, TITLE: Gradient-SDF: A Semi-Implicit Surface Representation for 3D Reconstruction
http://openaccess.thecvf.com//content/CVPR2022/html/Sommer_Gradient-SDF_A_Semi-Implicit_Surface_Representation_for_3D_Reconstruction_CVPR_2022_paper.html
AUTHORS: Christiane Sommer, Lu Sang, David Schubert, Daniel Cremers
HIGHLIGHT: We present Gradient-SDF, a novel representation for 3D geometry that combines the advantages of implict and explicit representations.

1847, TITLE: Bilateral Video Magnification Filter
http://openaccess.thecvf.com//content/CVPR2022/html/Takeda_Bilateral_Video_Magnification_Filter_CVPR_2022_paper.html
AUTHORS: Shoichiro Takeda, Kenta Niwa, Mariko Isogawa, Shinya Shimizu, Kazuki Okami, Yushi Aono

HIGHLIGHT: To enhance EVM performance on real videos, this paper proposes a bilateral video magnification filter (BVMF) that offers simple yet robust temporal filtering.

1848, TITLE: AdaFocus V2: End-to-End Training of Spatial Dynamic Networks for Video Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_AdaFocus_V2_End-to-End_Training_of_Spatial_Dynamic_Networks_for_Video_CVPR_2022_paper.html
AUTHORS: Yulin Wang, Yang Yue, Yuanze Lin, Haojun Jiang, Zihang Lai, Victor Kulikov, Nikita Orlov, Humphrey Shi, Gao Huang
HIGHLIGHT: This work reformulates the training of AdaFocus as a simple one-stage algorithm by introducing a differentiable interpolation-based patch selection operation, enabling efficient end-to-end optimization.

1849, TITLE: Localization Distillation for Dense Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Zheng_Localization_Distillation_for_Dense_Object_Detection_CVPR_2022_paper.html
AUTHORS: Zhaohui Zheng, Rongguang Ye, Ping Wang, Dongwei Ren, Wangmeng Zuo, Qibin Hou, Ming-Ming Cheng
HIGHLIGHT: In this paper, by reformulating the knowledge distillation process on localization, we present a novel localization distillation (LD) method which can efficiently transfer the localization knowledge from the teacher to the student.

1850, TITLE: What's in Your Hands? 3D Reconstruction of Generic Objects in Hands
http://openaccess.thecvf.com//content/CVPR2022/html/Ye_Whats_in_Your_Hands_3D_Reconstruction_of_Generic_Objects_in_CVPR_2022_paper.html
AUTHORS: Yufei Ye, Abhinav Gupta, Shubham Tulsiani
HIGHLIGHT: Our key insight is that hand articulation is highly predictive of the object shape, and we propose an approach that conditionally reconstructs the object based on the articulation and the visual input.

1851, TITLE: Continuous Scene Representations for Embodied AI
http://openaccess.thecvf.com//content/CVPR2022/html/Gadre_Continuous_Scene_Representations_for_Embodied_AI_CVPR_2022_paper.html
AUTHORS: Samir Yitzhak Gadre, Kiana Ehsani, Shuran Song, Roozbeh Mottaghi
HIGHLIGHT: We propose Continuous Scene Representations (CSR), a scene representation constructed by an embodied agent navigating within a space, where objects and their relationships are modeled by continuous valued embeddings.

1852, TITLE: Beyond 3D Siamese Tracking: A Motion-Centric Paradigm for 3D Single Object Tracking in Point Clouds
http://openaccess.thecvf.com//content/CVPR2022/html/Zheng_Beyond_3D_Siamese_Tracking_A_Motion-Centric_Paradigm_for_3D_Single_CVPR_2022_paper.html
AUTHORS: Chaoda Zheng, Xu Yan, Haiming Zhang, Baoyuan Wang, Shenghui Cheng, Shuguang Cui, Zhen Li
HIGHLIGHT: In this work, beyond 3D Siamese tracking, we introduce a motion-centric paradigm to handle 3D SOT from a new perspective.

1853, TITLE: Neural Mean Discrepancy for Efficient Out-of-Distribution Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Dong_Neural_Mean_Discrepancy_for_Efficient_Out-of-Distribution_Detection_CVPR_2022_paper.html
AUTHORS: Xin Dong, Junfeng Guo, Ang Li, Wei-Te Ting, Cong Liu, H.T. Kung
HIGHLIGHT: Based upon this observation, we propose a novel metric called Neural Mean Discrepancy (NMD), which compares neural means of the input examples and training data. Leveraging the simplicity of NMD, we propose an efficient OOD detector that computes neural means by a standard forward pass followed by a lightweight classifier.

1854, TITLE: Non-Probability Sampling Network for Stochastic Human Trajectory Prediction
http://openaccess.thecvf.com//content/CVPR2022/html/Bae_Non-Probability_Sampling_Network_for_Stochastic_Human_Trajectory_Prediction_CVPR_2022_paper.html
AUTHORS: Inhwan Bae, Jin-Hwi Park, Hae-Gon Jeon
HIGHLIGHT: In this paper, we analyze the problem by reconstructing and comparing probabilistic distributions from prediction samples and socially-acceptable paths, respectively.

1855, TITLE: Marginal Contrastive Correspondence for Guided Image Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Zhan_Marginal_Contrastive_Correspondence_for_Guided_Image_Generation_CVPR_2022_paper.html
AUTHORS: Fangneng Zhan, Yingchen Yu, Rongliang Wu, Jiahui Zhang, Shijian Lu, Changgong Zhang
HIGHLIGHT: We design a Marginal Contrastive Learning Network (MCL-Net) that explores contrastive learning to learn domain-invariant features for realistic exemplar-based image translation.

1856, TITLE: Complex Backdoor Detection by Symmetric Feature Differencing
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Complex_Backdoor_Detection_by_Symmetric_Feature_Differencing_CVPR_2022_paper.html
AUTHORS: Yingqi Liu, Guangyu Shen, Guanhong Tao, Zhenting Wang, Shiqing Ma, Xiangyu Zhang
HIGHLIGHT: We develop a new detection method. It first uses a trigger inversion technique to generate triggers, namely, universal input patterns flipping victim class samples to a target class. It then checks if any such trigger is composed of features that are not natural distinctive features between the victim and target classes.

1857, TITLE: Time Lens++: Event-Based Frame Interpolation With Parametric Non-Linear Flow and Multi-Scale Fusion

http://openaccess.thecvf.com//content/CVPR2022/html/Tulyakov_Time_Lens_Event-Based_Frame_Interpolation_With_Parametric_Non-Linear_Flow_and_CVPR_2022_paper.html
AUTHORS: Stepan Tulyakov, Alfredo Bochicchio, Daniel Gehrig, Stamatios Georgoulis, Yuanyou Li, Davide Scaramuzza
HIGHLIGHT: Moreover, previous methods were only tested on datasets consisting of planar and far-away scenes, which do not capture the full complexity of the real world. In this work, we address the above problems by introducing multi-scale feature-level fusion and computing one-shot non-linear inter-frame motion---which can be efficiently sampled for image warping---from events and images.

1858, TITLE: ResSFL: A Resistance Transfer Framework for Defending Model Inversion Attack in Split Federated Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Li_ResSFL_A_Resistance_Transfer_Framework_for_Defending_Model_Inversion_Attack_CVPR_2022_paper.html
AUTHORS: Jingtao Li, Adnan Siraj Rakin, Xing Chen, Zhezhi He, Deliang Fan, Chaitali Chakrabarti
HIGHLIGHT: So we propose ResSFL, a Split Federated Learning Framework that is designed to be MI-resistant during training.

1859, TITLE: RecDis-SNN: Rectifying Membrane Potential Distribution for Directly Training Spiking Neural Networks
http://openaccess.thecvf.com//content/CVPR2022/html/Guo_RecDis-SNN_Rectifying_Membrane_Potential_Distribution_for_Directly_Training_Spiking_Neural_CVPR_2022_paper.html
AUTHORS: Yufei Guo, Xinyi Tong, Yuanpei Chen, Liwen Zhang, Xiaode Liu, Zhe Ma, Xuhui Huang
HIGHLIGHT: Such undesired shifts would prevent the SNN from performing well and going deep. To tackle these problems, we attempt to rectify the membrane potential distribution (MPD) by designing a novel distribution loss, MPD-Loss, which can explicitly penalize the undesired shifts without introducing any additional operations in the inference phase.

1860, TITLE: Human-Aware Object Placement for Visual Environment Reconstruction
http://openaccess.thecvf.com//content/CVPR2022/html/Yi_Human-Aware_Object_Placement_for_Visual_Environment_Reconstruction_CVPR_2022_paper.html
AUTHORS: Hongwei Yi, Chun-Hao P. Huang, Dimitrios Tzionas, Muhammed Kocabas, Mohamed Hassan, Siyu Tang, Justus Thies, Michael J. Black
HIGHLIGHT: Our key idea is that, as a person moves through a scene and interacts with it, we accumulate HSIs across multiple input images, and use these in optimizing the 3D scene to reconstruct a consistent, physically plausible, 3D scene layout.

1861, TITLE: X-Pool: Cross-Modal Language-Video Attention for Text-Video Retrieval
http://openaccess.thecvf.com//content/CVPR2022/html/Gorti_X-Pool_Cross-Modal_Language-Video_Attention_for_Text-Video_Retrieval_CVPR_2022_paper.html
AUTHORS: Satya Krishna Gorti, Noël Vouitsis, Junwei Ma, Keyvan Golestan, Maksims Volkovs, Animesh Garg, Guangwei Yu
HIGHLIGHT: Common text-agnostic aggregations schemes include mean-pooling or self-attention over the frames, but these are likely to encode misleading visual information not described in the given text. To address this, we propose a cross-modal attention model called X-Pool that reasons between a text and the frames of a video.

1862, TITLE: Learning of Global Objective for Network Flow in Multi-Object Tracking
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Learning_of_Global_Objective_for_Network_Flow_in_Multi-Object_Tracking_CVPR_2022_paper.html
AUTHORS: Shuai Li, Yu Kong, Hamid Rezatofighi
HIGHLIGHT: Most previous studies focus on learning the cost function by only taking into account two frames during training, therefore the learned cost function is sub-optimal for MCF where a multi-frame data association must be considered during inference. In order to address this problem, in this paper we propose a novel differentiable framework that ties training and inference together during learning by solving a bi-level optimization problem, where the lower-level solves a linear program and the upper-level contains a loss function that incorporates global tracking result.

1863, TITLE: Towards Weakly-Supervised Text Spotting Using a Multi-Task Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Kittenplon_Towards_Weakly-Supervised_Text_Spotting_Using_a_Multi-Task_Transformer_CVPR_2022_paper.html
AUTHORS: Yair Kittenplon, Inbal Lavi, Sharon Fogel, Yarin Bar, R. Manmatha, Pietro Perona
HIGHLIGHT: We introduce TextTranSpotter (TTS), a transformer-based approach for text spotting and the first text spotting framework which may be trained with both fully- and weakly-supervised settings.

1864, TITLE: Gated2Gated: Self-Supervised Depth Estimation From Gated Images
http://openaccess.thecvf.com//content/CVPR2022/html/Walia_Gated2Gated_Self-Supervised_Depth_Estimation_From_Gated_Images_CVPR_2022_paper.html
AUTHORS: Amanpreet Walia, Stefanie Walz, Mario Bijelic, Fahim Mannan, Frank Julca-Aguilar, Michael Langer, Werner Ritter, Felix Heide
HIGHLIGHT: Although existing methods have shown that it is possible to decode high-resolution depth from such measurements, these methods require synchronized and calibrated LiDAR to supervise the gated depth decoder - prohibiting fast adoption across geographies, training on large unpaired datasets, and exploring alternative applications outside of automotive use cases. In this work, we fill this gap and propose an entirely self-supervised depth estimation method that uses gated intensity profiles and temporal consistency as a training signal.

1865, TITLE: RAMA: A Rapid Multicut Algorithm on GPU

http://openaccess.thecvf.com//content/CVPR2022/html/Abbas_RAMA_A_Rapid_Multicut_Algorithm_on_GPU_CVPR_2022_paper.html

AUTHORS: Ahmed Abbas, Paul Swoboda

HIGHLIGHT: We propose a highly parallel primal-dual algorithm for the multicut (a.k.a. correlation clustering) problem, a classical graph clustering problem widely used in machine learning and computer vision.

1866, TITLE: Adversarial Parametric Pose Prior
http://openaccess.thecvf.com//content/CVPR2022/html/Davydov_Adversarial_Parametric_Pose_Prior_CVPR_2022_paper.html

AUTHORS: Andrey Davydov, Anastasia Remizova, Victor Constantin, Sina Honari, Mathieu Salzmann, Pascal Fua

HIGHLIGHT: We propose learning a prior that restricts the SMPL parameters to values that produce realistic poses via adversarial training.

1867, TITLE: DC-SSL: Addressing Mismatched Class Distribution in Semi-Supervised Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Zhao_DC-SSL_Addressing_Mismatched_Class_Distribution_in_Semi-Supervised_Learning_CVPR_2022_paper.html

AUTHORS: Zhen Zhao, Luping Zhou, Yue Duan, Lei Wang, Lei Qi, Yinghuan Shi

HIGHLIGHT: To this end, this paper revisits the Exponentially Moving Average (EMA) model and utilizes it to estimate RCD in an iteratively improved manner, which is achieved with a momentum-update scheme throughout the training procedure.

1868, TITLE: Mask Transfiner for High-Quality Instance Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Ke_Mask_Transfiner_for_High-Quality_Instance_Segmentation_CVPR_2022_paper.html

AUTHORS: Lei Ke, Martin Danelljan, Xia Li, Yu-Wing Tai, Chi-Keung Tang, Fisher Yu

HIGHLIGHT: In this paper, we present Mask Transfiner for high-quality and efficient instance segmentation.

1869, TITLE: End-to-End Reconstruction-Classification Learning for Face Forgery Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Cao_End-to-End_Reconstruction-Classification_Learning_for_Face_Forgery_Detection_CVPR_2022_paper.html

AUTHORS: Junyi Cao, Chao Ma, Taiping Yao, Shen Chen, Shouhong Ding, Xiaokang Yang

HIGHLIGHT: In this paper, from a new perspective, we propose a forgery detection framework emphasizing the common compact representations of genuine faces based on reconstruction-classification learning.

1870, TITLE: It Is Okay To Not Be Okay: Overcoming Emotional Bias in Affective Image Captioning by Contrastive Data Collection
http://openaccess.thecvf.com//content/CVPR2022/html/Mohamed_It_Is_Okay_To_Not_Be_Okay_Overcoming_Emotional_Bias_CVPR_2022_paper.html

AUTHORS: Youssef Mohamed, Faizan Farooq Khan, Kilichbek Haydarov, Mohamed Elhoseiny

HIGHLIGHT: We show that collecting new data, in the same way, is not effective in mitigating this emotional bias. To remedy this problem, we propose a contrastive data collection approach to balance ArtEmis with a new complementary dataset such that a pair of similar images have contrasting emotions (one positive and one negative).

1871, TITLE: Transferability Metrics for Selecting Source Model Ensembles
http://openaccess.thecvf.com//content/CVPR2022/html/Agostinelli_Transferability_Metrics_for_Selecting_Source_Model_Ensembles_CVPR_2022_paper.html

AUTHORS: Andrea Agostinelli, Jasper Uijlings, Thomas Mensink, Vittorio Ferrari

HIGHLIGHT: Since fine-tuning all possible ensembles is computationally prohibitive, we aim at predicting performance on the target dataset using a computationally efficient transferability metric.

1872, TITLE: Neural Global Shutter: Learn To Restore Video From a Rolling Shutter Camera With Global Reset Feature
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Neural_Global_Shutter_Learn_To_Restore_Video_From_a_Rolling_CVPR_2022_paper.html

AUTHORS: Zhixiang Wang, Xiang Ji, Jia-Bin Huang, Shin'ichi Satoh, Xiao Zhou, Yinqiang Zheng

HIGHLIGHT: In this paper, we investigate using rolling shutter with a global reset feature (RSGR) to restore clean global shutter (GS) videos.

1873, TITLE: DiRA: Discriminative, Restorative, and Adversarial Learning for Self-Supervised Medical Image Analysis
http://openaccess.thecvf.com//content/CVPR2022/html/Haghighi_DiRA_Discriminative_Restorative_and_Adversarial_Learning_for_Self-Supervised_Medical_Image_CVPR_2022_paper.html

AUTHORS: Fatemeh Haghighi, Mohammad Reza Hosseinzadeh Taher, Michael B. Gotway, Jianming Liang

HIGHLIGHT: Existing efforts, however, omit their synergistic effects on each other in a ternary setup, which, we envision, can significantly benefit deep semantic representation learning. To realize this vision, we have developed DiRA, the first framework that unites discriminative, restorative, and adversarial learning in a unified manner to collaboratively glean complementary visual information from unlabeled medical images for fine-grained semantic representation learning.

1874, TITLE: Open Challenges in Deep Stereo: The Booster Dataset
http://openaccess.thecvf.com//content/CVPR2022/html/Ramirez_Open_Challenges_in_Deep_Stereo_The_Booster_Dataset_CVPR_2022_paper.html

AUTHORS: Pierluigi Zama Ramirez, Fabio Tosi, Matteo Poggi, Samuele Salti, Stefano Mattoccia, Luigi Di Stefano

HIGHLIGHT: We present a novel high-resolution and challenging stereo dataset framing indoor scenes annotated with dense and accurate ground-truth disparities.

1875, TITLE:	Location-Free Human Pose Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_Location-Free_Human_Pose_Estimation_CVPR_2022_paper.html
AUTHORS:	Xixia Xu, Yingguo Gao, Ke Yan, Xue Lin, Qi Zou
HIGHLIGHT:	However, it is rather time-consuming to collect high-quality and fine-grained annotations for human body. To alleviate this issue, we revisit HPE and propose a location-free framework without supervision of keypoint locations.

1876, TITLE:	Self-Supervised Bulk Motion Artifact Removal in Optical Coherence Tomography Angiography
http://openaccess.thecvf.com//content/CVPR2022/html/Ren_Self-Supervised_Bulk_Motion_Artifact_Removal_in_Optical_Coherence_Tomography_Angiography_CVPR_2022_paper.html
AUTHORS:	Jiaxiang Ren, Kicheon Park, Yingtian Pan, Haibin Ling
HIGHLIGHT:	In addition, these methods discard the rich structural and appearance information carried in the BMA stripe region. To address these issues, in this paper we propose a self-supervised content-aware BMA removal model.

1877, TITLE:	Watch It Move: Unsupervised Discovery of 3D Joints for Re-Posing of Articulated Objects
http://openaccess.thecvf.com//content/CVPR2022/html/Noguchi_Watch_It_Move_Unsupervised_Discovery_of_3D_Joints_for_Re-Posing_CVPR_2022_paper.html
AUTHORS:	Atsuhiro Noguchi, Umar Iqbal, Jonathan Tremblay, Tatsuya Harada, Orazio Gallo
HIGHLIGHT:	Unfortunately, assuming the structure to be known, as existing methods do, precludes the ability to work on new object categories. We propose to learn both the appearance and the structure of previously unseen articulated objects by observing them move from multiple views, with no joints annotation supervision, or information about the structure.

1878, TITLE:	PoseTrack21: A Dataset for Person Search, Multi-Object Tracking and Multi-Person Pose Tracking
http://openaccess.thecvf.com//content/CVPR2022/html/Doring_PoseTrack21_A_Dataset_for_Person_Search_Multi-Object_Tracking_and_Multi-Person_CVPR_2022_paper.html
AUTHORS:	Andreas Döring, Di Chen, Shanshan Zhang, Bernt Schiele, Jürgen Gall
HIGHLIGHT:	Consequently, approaches on these respective tasks are eligible to complement each other. Therefore, we introduce PoseTrack21, a large-scale dataset for person search, multi-object tracking and multi-person pose tracking in real-world scenarios with a high diversity of poses.

1879, TITLE:	Event-Based Video Reconstruction via Potential-Assisted Spiking Neural Network
http://openaccess.thecvf.com//content/CVPR2022/html/Zhu_Event-Based_Video_Reconstruction_via_Potential-Assisted_Spiking_Neural_Network_CVPR_2022_paper.html
AUTHORS:	Lin Zhu, Xiao Wang, Yi Chang, Jianing Li, Tiejun Huang, Yonghong Tian
HIGHLIGHT:	We propose a novel Event-based Video reconstruction framework based on a fully Spiking Neural Network (EVSNN), which utilizes Leaky-Integrate-and-Fire (LIF) neuron and Membrane Potential (MP) neuron.

1880, TITLE:	Efficient Maximal Coding Rate Reduction by Variational Forms
http://openaccess.thecvf.com//content/CVPR2022/html/Baek_Efficient_Maximal_Coding_Rate_Reduction_by_Variational_Forms_CVPR_2022_paper.html
AUTHORS:	Christina Baek, Ziyang Wu, Kwan Ho Ryan Chan, Tianjiao Ding, Yi Ma, Benjamin D. Haeffele
HIGHLIGHT:	By taking advantage of variational forms of spectral functions of a matrix, we reformulate the MCR2 objective to a form that can scale significantly without compromising training accuracy.

1881, TITLE:	Ithaca365: Dataset and Driving Perception Under Repeated and Challenging Weather Conditions
http://openaccess.thecvf.com//content/CVPR2022/html/Diaz-Ruiz_Ithaca365_Dataset_and_Driving_Perception_Under_Repeated_and_Challenging_Weather_CVPR_2022_paper.html
AUTHORS:	Carlos A. Diaz-Ruiz, Youya Xia, Yurong You, Jose Nino, Junan Chen, Josephine Monica, Xiangyu Chen, Katie Luo, Yan Wang, Marc Emond, Wei-Lun Chao, Bharath Hariharan, Kilian Q. Weinberger, Mark Campbell
HIGHLIGHT:	In this paper, we present a new dataset to enable robust autonomous driving via a novel data collection process --- data is repeatedly recorded along a 15 km route under diverse scene (urban, highway, rural, campus), weather (snow, rain, sun), time (day/night), and traffic conditions (pedestrians, cyclists and cars).

1882, TITLE:	AutoLoss-GMS: Searching Generalized Margin-Based Softmax Loss Function for Person Re-Identification
http://openaccess.thecvf.com//content/CVPR2022/html/Gu_AutoLoss-GMS_Searching_Generalized_Margin-Based_Softmax_Loss_Function_for_Person_Re-Identification_CVPR_2022_paper.html
AUTHORS:	Hongyang Gu, Jianmin Li, Guangyuan Fu, Chifong Wong, Xinghao Chen, Jun Zhu
HIGHLIGHT:	In this paper, we propose a novel method, AutoLoss-GMS, to search the better loss function in the space of generalized margin-based softmax loss function for person re-identification automatically.

1883, TITLE:	YouMVOS: An Actor-Centric Multi-Shot Video Object Segmentation Dataset
http://openaccess.thecvf.com//content/CVPR2022/html/Wei_YouMVOS_An_Actor-Centric_Multi-Shot_Video_Object_Segmentation_Dataset_CVPR_2022_paper.html
AUTHORS:	Donglai Wei, Siddhant Kharbanda, Sarthak Arora, Roshan Roy, Nishant Jain, Akash Palrecha, Tanav Shah, Shray Mathur, Ritik Mathur, Abhijay Kemkar, Anirudh Chakravarthy, Zudi Lin, Won-Dong Jang, Yansong Tang, Song Bai, James Tompkin, Philip H.S. Torr, Hanspeter Pfister
HIGHLIGHT:	Many video understanding tasks require analyzing multi-shot videos, but existing datasets for video object segmentation (VOS) only consider single-shot videos. To address this challenge, we collected a new dataset---YouMVOS---of 200 popular YouTube videos spanning ten genres, where each video is on average five minutes long and with 75 shots.

1884, TITLE: DAFormer: Improving Network Architectures and Training Strategies for Domain-Adaptive Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Hoyer_DAFormer_Improving_Network_Architectures_and_Training_Strategies_for_Domain-Adaptive_Semantic_CVPR_2022_paper.html
AUTHORS: Lukas Hoyer, Dengxin Dai, Luc Van Gool
HIGHLIGHT: As the influence of recent network architectures has not been systematically studied, we first benchmark different network architectures for UDA and newly reveal the potential of Transformers for UDA semantic segmentation. Based on the findings, we propose a novel UDA method, DAFormer.

1885, TITLE: Sound-Guided Semantic Image Manipulation
http://openaccess.thecvf.com//content/CVPR2022/html/Lee_Sound-Guided_Semantic_Image_Manipulation_CVPR_2022_paper.html
AUTHORS: Seung Hyun Lee, Wonseok Roh, Wonmin Byeon, Sang Ho Yoon, Chanyoung Kim, Jinkyu Kim, Sangpil Kim
HIGHLIGHT: Here, we propose a framework that directly encodes sound into the multi-modal (image-text) embedding space and manipulates an image from the space.

1886, TITLE: Joint Distribution Matters: Deep Brownian Distance Covariance for Few-Shot Classification
http://openaccess.thecvf.com//content/CVPR2022/html/Xie_Joint_Distribution_Matters_Deep_Brownian_Distance_Covariance_for_Few-Shot_Classification_CVPR_2022_paper.html
AUTHORS: Jiangtao Xie, Fei Long, Jiaming Lv, Qilong Wang, Peihua Li
HIGHLIGHT: In this paper, we propose a deep Brownian Distance Covariance (DeepBDC) method for few-shot classification.

1887, TITLE: Proper Reuse of Image Classification Features Improves Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Vasconcelos_Proper_Reuse_of_Image_Classification_Features_Improves_Object_Detection_CVPR_2022_paper.html
AUTHORS: Cristina Vasconcelos, Vighnesh Birodkar, Vincent Dumoulin
HIGHLIGHT: Recent works show this is not strictly necessary under longer training regimes and provide recipes for training the backbone from scratch. We investigate the opposite direction of this end-to-end training trend: we show that an extreme form of knowledge preservation--freezing the classifier-initialized backbone-- consistently improves many different detection models, and leads to considerable resource savings.

1888, TITLE: MetaPose: Fast 3D Pose From Multiple Views Without 3D Supervision
http://openaccess.thecvf.com//content/CVPR2022/html/Usman_MetaPose_Fast_3D_Pose_From_Multiple_Views_Without_3D_Supervision_CVPR_2022_paper.html
AUTHORS: Ben Usman, Andrea Tagliasacchi, Kate Saenko, Avneesh Sud
HIGHLIGHT: In the era of deep learning, human pose estimation from multiple cameras with unknown calibration has received little attention to date. We show how to train a neural model to perform this task with high precision and minimal latency overhead.

1889, TITLE: End-to-End Human-Gaze-Target Detection With Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Tu_End-to-End_Human-Gaze-Target_Detection_With_Transformers_CVPR_2022_paper.html
AUTHORS: Danyang Tu, Xiongkuo Min, Huiyu Duan, Guodong Guo, Guangtao Zhai, Wei Shen
HIGHLIGHT: In this paper, we propose an effective and efficient method for Human-Gaze-Target (HGT) detection, i.e., gaze following.

1890, TITLE: The Devil Is in the Pose: Ambiguity-Free 3D Rotation-Invariant Learning via Pose-Aware Convolution
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_The_Devil_Is_in_the_Pose_Ambiguity-Free_3D_Rotation-Invariant_Learning_CVPR_2022_paper.html
AUTHORS: Ronghan Chen, Yang Cong
HIGHLIGHT: In this paper, we real that the global information loss stems from an unexplored pose information loss problem, i.e., common convolution layers cannot capture the relative poses between RI features, thus hindering the global information to be hierarchically aggregated in the deep networks.

1891, TITLE: Compositional Temporal Grounding With Structured Variational Cross-Graph Correspondence Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Compositional_Temporal_Grounding_With_Structured_Variational_Cross-Graph_Correspondence_Learning_CVPR_2022_paper.html
AUTHORS: Juncheng Li, Junlin Xie, Long Qian, Linchao Zhu, Siliang Tang, Fei Wu, Yi Yang, Yueting Zhuang, Xin Eric Wang
HIGHLIGHT: Evaluating the state-of-the-art methods on our new dataset splits, we empirically find that they fail to generalize to queries with novel combinations of seen words. To tackle this challenge, we propose a variational cross-graph reasoning framework that explicitly decomposes video and language into multiple structured hierarchies and learns fine-grained semantic correspondence among them.

1892, TITLE: Visible-Thermal UAV Tracking: A Large-Scale Benchmark and New Baseline
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Visible-Thermal_UAV_Tracking_A_Large-Scale_Benchmark_and_New_Baseline_CVPR_2022_paper.html
AUTHORS: Pengyu Zhang, Jie Zhao, Dong Wang, Huchuan Lu, Xiang Ruan
HIGHLIGHT: In this paper, we construct a large-scale benchmark with high diversity for visible-thermal UAV tracking (VTUAV), including 500 sequences with 1.7 million high-resolution (1920*1080 pixels) frame pairs.

1893, TITLE:	Future Transformer for Long-Term Action Anticipation
http://openaccess.thecvf.com//content/CVPR2022/html/Gong_Future_Transformer_for_Long-Term_Action_Anticipation_CVPR_2022_paper.html
AUTHORS:	Dayoung Gong, Joonseok Lee, Manjin Kim, Seong Jong Ha, Minsu Cho
HIGHLIGHT:	In a similar spirit, we propose an end-to-end attention model for action anticipation, dubbed Future Transformer (FUTR), that leverages global attention over all input frames and output tokens to predict a minutes-long sequence of future actions.

1894, TITLE:	Optimal LED Spectral Multiplexing for NIR2RGB Translation
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Optimal_LED_Spectral_Multiplexing_for_NIR2RGB_Translation_CVPR_2022_paper.html
AUTHORS:	Lei Liu, Yuze Chen, Junchi Yan, Yinqiang Zheng
HIGHLIGHT:	In this paper, we aim to break the fundamental restrictions on reliable NIR-to-RGB (NIR2RGB) translation by examining the imaging mechanism of single-chip silicon-based RGB cameras under NIR illuminations, and propose to retrieve the optimal LED multiplexing via deep learning.

1895, TITLE:	Rethinking Spatial Invariance of Convolutional Networks for Object Counting
http://openaccess.thecvf.com//content/CVPR2022/html/Cheng_Rethinking_Spatial_Invariance_of_Convolutional_Networks_for_Object_Counting_CVPR_2022_paper.html
AUTHORS:	Zhi-Qi Cheng, Qi Dai, Hong Li, Jingkuan Song, Xiao Wu, Alexander G. Hauptmann
HIGHLIGHT:	In this paper, we try to use locally connected Gaussian kernels to replace the original convolution filter to estimate the spatial position in the density map.

1896, TITLE:	Self-Supervised Video Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Ranasinghe_Self-Supervised_Video_Transformer_CVPR_2022_paper.html
AUTHORS:	Kanchana Ranasinghe, Muzammal Naseer, Salman Khan, Fahad Shahbaz Khan, Michael S. Ryoo
HIGHLIGHT:	In this paper, we propose self-supervised training for video transformers using unlabeled video data.

1897, TITLE:	AutoRF: Learning 3D Object Radiance Fields From Single View Observations
http://openaccess.thecvf.com//content/CVPR2022/html/Muller_AutoRF_Learning_3D_Object_Radiance_Fields_From_Single_View_Observations_CVPR_2022_paper.html
AUTHORS:	Norman Müller, Andrea Simonelli, Lorenzo Porzi, Samuel Rota Bulò, Matthias Nießner, Peter Kontschieder
HIGHLIGHT:	We introduce AutoRF - a new approach for learning neural 3D object representations where each object in the training set is observed by only a single view.

1898, TITLE:	Expanding Large Pre-Trained Unimodal Models With Multimodal Information Injection for Image-Text Multimodal Classification
http://openaccess.thecvf.com//content/CVPR2022/html/Liang_Expanding_Large_Pre-Trained_Unimodal_Models_With_Multimodal_Information_Injection_for_CVPR_2022_paper.html
AUTHORS:	Tao Liang, Guosheng Lin, Mingyang Wan, Tianrui Li, Guojun Ma, Fengmao Lv
HIGHLIGHT:	To this end, we propose the Multimodal Information Injection Plug-in (MI2P) which is attached to different layers of the unimodal models (e.g., DenseNet and BERT).

1899, TITLE:	Neural RGB-D Surface Reconstruction
http://openaccess.thecvf.com//content/CVPR2022/html/Azinovic_Neural_RGB-D_Surface_Reconstruction_CVPR_2022_paper.html
AUTHORS:	Dejan Azinović, Ricardo Martin-Brualla, Dan B Goldman, Matthias Nießner, Justus Thies
HIGHLIGHT:	The volumetric representation of the surface based on densities leads to artifacts when a surface is extracted using Marching Cubes, since during optimization, densities are accumulated along the ray and are not used at a single sample point in isolation. Instead of this volumetric representation of the surface, we propose to represent the surface using an implicit function (truncated signed distance function).

1900, TITLE:	ClusterGNN: Cluster-Based Coarse-To-Fine Graph Neural Network for Efficient Feature Matching
http://openaccess.thecvf.com//content/CVPR2022/html/Shi_ClusterGNN_Cluster-Based_Coarse-To-Fine_Graph_Neural_Network_for_Efficient_Feature_Matching_CVPR_2022_paper.html
AUTHORS:	Yan Shi, Jun-Xiong Cai, Yoli Shavit, Tai-Jiang Mu, Wensen Feng, Kai Zhang
HIGHLIGHT:	Motivated by a prior observation that self- and cross- attention matrices converge to a sparse representation, we propose ClusterGNN, an attentional GNN architecture which operates on clusters for learning the feature matching task.

1901, TITLE:	AdaptPose: Cross-Dataset Adaptation for 3D Human Pose Estimation by Learnable Motion Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Gholami_AdaptPose_Cross-Dataset_Adaptation_for_3D_Human_Pose_Estimation_by_Learnable_CVPR_2022_paper.html
AUTHORS:	Mohsen Gholami, Bastian Wandt, Helge Rhodin, Rabab Ward, Z. Jane Wang
HIGHLIGHT:	To this end, we propose AdaptPose, an end-to-end framework that generates synthetic 3D human motions from a source dataset and uses them to fine-tune a 3D pose estimator.

1902, TITLE:	ClothFormer: Taming Video Virtual Try-On in All Module
http://openaccess.thecvf.com//content/CVPR2022/html/Jiang_ClothFormer_Taming_Video_Virtual_Try-On_in_All_Module_CVPR_2022_paper.html
AUTHORS:	Jianbin Jiang, Tan Wang, He Yan, Junhui Liu

HIGHLIGHT: Moreover, there are two other key challenges: 1) how to generate accurate warping when occlusions appear in the clothing region; 2) how to generate clothes and non-target body parts (e.g. arms, neck) in harmony with the complicated background; To address them, we propose a novel video virtual try-on framework, ClothFormer, which successfully synthesizes realistic, harmonious, and spatio-temporal consistent results in complicated environment.

1903, TITLE: Cross-Domain Adaptive Teacher for Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Cross-Domain_Adaptive_Teacher_for_Object_Detection_CVPR_2022_paper.html
AUTHORS: Yu-Jhe Li, Xiaoliang Dai, Chih-Yao Ma, Yen-Cheng Liu, Kan Chen, Bichen Wu, Zijian He, Kris Kitani, Peter Vajda
HIGHLIGHT: However, it suffers from the domain shift and generates many low-quality pseudo labels (e.g., false positives), which leads to sub-optimal performance. To mitigate this problem, we propose a teacher-student framework named Adaptive Teacher (AT) which leverages domain adversarial learning and weak-strong data augmentation to address the domain gap.

1904, TITLE: Geometric Anchor Correspondence Mining With Uncertainty Modeling for Universal Domain Adaptation
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Geometric_Anchor_Correspondence_Mining_With_Uncertainty_Modeling_for_Universal_Domain_CVPR_2022_paper.html
AUTHORS: Liang Chen, Yihang Lou, Jianzhong He, Tao Bai, Minghua Deng
HIGHLIGHT: Previous works rarely explore the intrinsic geometrical relationship between the two domains, and they manually set a threshold for the overconfident closed-world classifier to reject "unknown" samples. Therefore, in this paper, we propose a Geometric anchor-guided Adversarial and conTrastive learning framework with uncErtainty modeling called GATE to alleviate these issues.

1905, TITLE: Class-Balanced Pixel-Level Self-Labeling for Domain Adaptive Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Class-Balanced_Pixel-Level_Self-Labeling_for_Domain_Adaptive_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Ruihuang Li, Shuai Li, Chenhang He, Yabin Zhang, Xu Jia, Lei Zhang
HIGHLIGHT: However, the produced pseudo labels often contain much noise because the model is biased to source domain as well as majority categories. To address the above issues, we propose to directly explore the intrinsic pixel distributions of target domain data, instead of heavily relying on the source domain.

1906, TITLE: Coopernaut: End-to-End Driving With Cooperative Perception for Networked Vehicles
http://openaccess.thecvf.com//content/CVPR2022/html/Cui_Coopernaut_End-to-End_Driving_With_Cooperative_Perception_for_Networked_Vehicles_CVPR_2022_paper.html
AUTHORS: Jiaxun Cui, Hang Qiu, Dian Chen, Peter Stone, Yuke Zhu
HIGHLIGHT: We introduce COOPERNAUT, an end-to-end learning model that uses cross-vehicle perception for vision-based cooperative driving.

1907, TITLE: Condensing CNNs With Partial Differential Equations
http://openaccess.thecvf.com//content/CVPR2022/html/Kag_Condensing_CNNs_With_Partial_Differential_Equations_CVPR_2022_paper.html
AUTHORS: Anil Kag, Venkatesh Saligrama
HIGHLIGHT: We propose a new feature layer, called the Global layer, that enforces PDE constraints on the feature maps, resulting in rich features.

1908, TITLE: Few-Shot Keypoint Detection With Uncertainty Learning for Unseen Species
http://openaccess.thecvf.com//content/CVPR2022/html/Lu_Few-Shot_Keypoint_Detection_With_Uncertainty_Learning_for_Unseen_Species_CVPR_2022_paper.html
AUTHORS: Changsheng Lu, Piotr Koniusz
HIGHLIGHT: Thus, we propose a versatile Few-shot Keypoint Detection (FSKD) pipeline, which can detect a varying number of keypoints of different kinds.

1909, TITLE: Improving Robustness Against Stealthy Weight Bit-Flip Attacks by Output Code Matching
http://openaccess.thecvf.com//content/CVPR2022/html/Ozdenizci_Improving_Robustness_Against_Stealthy_Weight_Bit-Flip_Attacks_by_Output_Code_CVPR_2022_paper.html
AUTHORS: Ozan Özdenizci, Robert Legenstein
HIGHLIGHT: This renders the attack undetectable from a DNN operation perspective. We propose a DNN defense mechanism to improve robustness in such realistic stealthy weight bit-flip attack scenarios.

1910, TITLE: Unsupervised Hierarchical Semantic Segmentation With Multiview Cosegmentation and Clustering Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Ke_Unsupervised_Hierarchical_Semantic_Segmentation_With_Multiview_Cosegmentation_and_Clustering_Transformers_CVPR_2022_paper.html
AUTHORS: Tsung-Wei Ke, Jyh-Jing Hwang, Yunhui Guo, Xudong Wang, Stella X. Yu
HIGHLIGHT: Our idea is that a good representation must be able to reveal not just a particular level of grouping, but any level of grouping in a consistent and predictable manner across different levels of granularity.

1911, TITLE: 3D-SPS: Single-Stage 3D Visual Grounding via Referred Point Progressive Selection
http://openaccess.thecvf.com//content/CVPR2022/html/Luo_3D-SPS_Single-Stage_3D_Visual_Grounding_via_Referred_Point_Progressive_Selection_CVPR_2022_paper.html

AUTHORS: Junyu Luo, Jiahui Fu, Xianghao Kong, Chen Gao, Haibing Ren, Hao Shen, Huaxia Xia, Si Liu
HIGHLIGHT: In this paper, we propose a 3D Single-Stage Referred Point Progressive Selection (3D-SPS) method, which progressively selects keypoints with the guidance of language and directly locates the target.

1912, TITLE: TubeR: Tubelet Transformer for Video Action Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Zhao_TubeR_Tubelet_Transformer_for_Video_Action_Detection_CVPR_2022_paper.html
AUTHORS: Jiaojiao Zhao, Yanyi Zhang, Xinyu Li, Hao Chen, Bing Shuai, Mingze Xu, Chunhui Liu, Kaustav Kundu, Yuanjun Xiong, Davide Modolo, Ivan Marsic, Cees G. M. Snoek, Joseph Tighe
HIGHLIGHT: We propose TubeR: a simple solution for spatio-temporal video action detection.

1913, TITLE: LASER: LAtent SpacE Rendering for 2D Visual Localization
http://openaccess.thecvf.com//content/CVPR2022/html/Min_LASER_LAtent_SpacE_Rendering_for_2D_Visual_Localization_CVPR_2022_paper.html
AUTHORS: Zhixiang Min, Naji Khosravan, Zachary Bessinger, Manjunath Narayana, Sing Bing Kang, Enrique Dunn, Ivaylo Boyadzhiev
HIGHLIGHT: We present LASER, an image-based Monte Carlo Localization (MCL) framework for 2D floor maps.

1914, TITLE: MUM: Mix Image Tiles and UnMix Feature Tiles for Semi-Supervised Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Kim_MUM_Mix_Image_Tiles_and_UnMix_Feature_Tiles_for_Semi-Supervised_CVPR_2022_paper.html
AUTHORS: JongMok Kim, JooYoung Jang, Seunghyeon Seo, Jisoo Jeong, Jongkeun Na, Nojun Kwak
HIGHLIGHT: Especially when extending SSL to semi-supervised object detection (SSOD), many strong augmentation methodologies related to image geometry and interpolation-regularization are hard to utilize since they possibly hurt the location information of the bounding box in the object detection task. To address this, we introduce a simple yet effective data augmentation method, Mix/UnMix (MUM), which unmixes feature tiles for the mixed image tiles for the SSOD framework.

1915, TITLE: On Adversarial Robustness of Trajectory Prediction for Autonomous Vehicles
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_On_Adversarial_Robustness_of_Trajectory_Prediction_for_Autonomous_Vehicles_CVPR_2022_paper.html
AUTHORS: Qingzhao Zhang, Shengtuo Hu, Jiachen Sun, Qi Alfred Chen, Z. Morley Mao
HIGHLIGHT: However, few studies have analyzed the adversarial robustness of trajectory prediction or investigated whether the worst-case prediction can still lead to safe planning. To bridge this gap, we study the adversarial robustness of trajectory prediction models by proposing a new adversarial attack that perturbs normal vehicle trajectories to maximize the prediction error.

1916, TITLE: Kubric: A Scalable Dataset Generator
http://openaccess.thecvf.com//content/CVPR2022/html/Greff_Kubric_A_Scalable_Dataset_Generator_CVPR_2022_paper.html
AUTHORS: Klaus Greff, Francois Belletti, Lucas Beyer, Carl Doersch, Yilun Du, Daniel Duckworth, David J. Fleet, Dan Gnanapragasam, Florian Golemo, Charles Herrmann, Thomas Kipf, Abhijit Kundu, Dmitry Lagun, Issam Laradji, Hsueh-Ti (Derek) Liu, Henning Meyer, Yishu Miao, Derek Nowrouzezahrai, Cengiz Oztireli, Etienne Pot, Noha Radwan, Daniel Rebain, Sara Sabour, Mehdi S. M. Sajjadi, Matan Sela, Vincent Sitzmann, Austin Stone, Deqing Sun, Suhani Vora, Ziyu Wang, Tianhao Wu, Kwang Moo Yi, Fangcheng Zhong, Andrea Tagliasacchi
HIGHLIGHT: Unfortunately, software tools for effective data generation are less mature than those for architecture design and training, which leads to fragmented generation efforts. To address these problems we introduce Kubric, an open-source Python framework that interfaces with PyBullet and Blender to generate photo-realistic scenes, with rich annotations, and seamlessly scales to large jobs distributed over thousands of machines, and generating TBs of data.

1917, TITLE: Unpaired Deep Image Deraining Using Dual Contrastive Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Unpaired_Deep_Image_Deraining_Using_Dual_Contrastive_Learning_CVPR_2022_paper.html
AUTHORS: Xiang Chen, Jinshan Pan, Kui Jiang, Yufeng Li, Yufeng Huang, Caihua Kong, Longgang Dai, Zhentao Fan
HIGHLIGHT: In this paper, we develop an effective unpaired SID adversarial framework which explores mutual properties of the unpaired exemplars by a dual contrastive learning manner in a deep feature space, named as DCD-GAN.

1918, TITLE: Learning Multiple Dense Prediction Tasks From Partially Annotated Data
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Learning_Multiple_Dense_Prediction_Tasks_From_Partially_Annotated_Data_CVPR_2022_paper.html
AUTHORS: Wei-Hong Li, Xialei Liu, Hakan Bilen
HIGHLIGHT: In this paper, we present a label efficient approach and look at jointly learning of multiple dense prediction tasks on partially annotated data (i.e. not all the task labels are available for each image), which we call multi-task partially-supervised learning.

1919, TITLE: Pushing the Performance Limit of Scene Text Recognizer Without Human Annotation
http://openaccess.thecvf.com//content/CVPR2022/html/Zheng_Pushing_the_Performance_Limit_of_Scene_Text_Recognizer_Without_Human_CVPR_2022_paper.html
AUTHORS: Caiyuan Zheng, Hui Li, Seon-Min Rhee, Seungju Han, Jae-Joon Han, Peng Wang
HIGHLIGHT: In this work, we aim to boost STR models by leveraging both synthetic data and the numerous real unlabeled images, exempting human annotation cost thoroughly.

1920, TITLE: Boosting 3D Object Detection by Simulating Multimodality on Point Clouds

http://openaccess.thecvf.com//content/CVPR2022/html/Zheng_Boosting_3D_Object_Detection_by_Simulating_Multimodality_on_Point_Clouds_CVPR_2022_paper.html
AUTHORS: Wu Zheng, Mingxuan Hong, Li Jiang, Chi-Wing Fu
HIGHLIGHT: This paper presents a new approach to boost a single-modality (LiDAR) 3D object detector by teaching it to simulate features and responses that follow a multi-modality (LiDAR-image) detector.

1921, TITLE: Towards Low-Cost and Efficient Malaria Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Sultani_Towards_Low-Cost_and_Efficient_Malaria_Detection_CVPR_2022_paper.html
AUTHORS: Waqas Sultani, Wajahat Nawaz, Syed Javed, Muhammad Sohail Danish, Asma Saadia, Mohsen Ali
HIGHLIGHT: We present a dataset to further the research on malaria microscopy over the low-cost microscopes at low magnification.

1922, TITLE: Learning Neural Light Fields With Ray-Space Embedding
http://openaccess.thecvf.com//content/CVPR2022/html/Attal_Learning_Neural_Light_Fields_With_Ray-Space_Embedding_CVPR_2022_paper.html
AUTHORS: Benjamin Attal, Jia-Bin Huang, Michael Zollhöfer, Johannes Kopf, Changil Kim
HIGHLIGHT: We present a novel neural light field representation that, in contrast to prior work, is fast, memory efficient, and excels at modeling complicated view dependence.

1923, TITLE: Exposure Normalization and Compensation for Multiple-Exposure Correction
http://openaccess.thecvf.com//content/CVPR2022/html/Huang_Exposure_Normalization_and_Compensation_for_Multiple-Exposure_Correction_CVPR_2022_paper.html
AUTHORS: Jie Huang, Yajing Liu, Xueyang Fu, Man Zhou, Yang Wang, Feng Zhao, Zhiwei Xiong
HIGHLIGHT: Besides, to further alleviate the imbalanced performance caused by variations in the optimization process, we introduce a parameter regularization fine-tuning strategy to improve the performance of the worst-performed exposure without degrading other exposures.

1924, TITLE: UDA-COPE: Unsupervised Domain Adaptation for Category-Level Object Pose Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Lee_UDA-COPE_Unsupervised_Domain_Adaptation_for_Category-Level_Object_Pose_Estimation_CVPR_2022_paper.html
AUTHORS: Taeyeop Lee, Byeong-Uk Lee, Inkyu Shin, Jaesung Choe, Ukcheol Shin, In So Kweon, Kuk-Jin Yoon
HIGHLIGHT: Learning to estimate object pose often requires ground-truth (GT) labels, such as CAD model and absolute-scale object pose, which is expensive and laborious to obtain in the real world. To tackle this problem, we propose an unsupervised domain adaptation (UDA) for category-level object pose estimation, called UDA-COPE.

1925, TITLE: Learning Non-Target Knowledge for Few-Shot Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Learning_Non-Target_Knowledge_for_Few-Shot_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Yuanwei Liu, Nian Liu, Qinglong Cao, Xiwen Yao, Junwei Han, Ling Shao
HIGHLIGHT: Existing studies in few-shot semantic segmentation only focus on mining the target object information, however, often are hard to tell ambiguous regions, especially in non-target regions, which include background (BG) and Distracting Objects (DOs). To alleviate this problem, we propose a novel framework, namely Non-Target Region Eliminating (NTRE) network, to explicitly mine and eliminate BG and DO regions in the query.

1926, TITLE: TransFusion: Robust LiDAR-Camera Fusion for 3D Object Detection With Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Bai_TransFusion_Robust_LiDAR-Camera_Fusion_for_3D_Object_Detection_With_Transformers_CVPR_2022_paper.html
AUTHORS: Xuyang Bai, Zeyu Hu, Xinge Zhu, Qingqiu Huang, Yilun Chen, Hongbo Fu, Chiew-Lan Tai
HIGHLIGHT: We propose TransFusion, a robust solution to LiDAR-camera fusion with a soft-association mechanism to handle inferior image conditions.

1927, TITLE: Real-Time Hyperspectral Imaging in Hardware via Trained Metasurface Encoders
http://openaccess.thecvf.com//content/CVPR2022/html/Makarenko_Real-Time_Hyperspectral_Imaging_in_Hardware_via_Trained_Metasurface_Encoders_CVPR_2022_paper.html
AUTHORS: Maksim Makarenko, Arturo Burguete-Lopez, Qizhou Wang, Fedor Getman, Silvio Giancola, Bernard Ghanem, Andrea Fratalocchi
HIGHLIGHT: This work introduces Hyplex, a new integrated architecture addressing the limitations discussed above.

1928, TITLE: Clean Implicit 3D Structure From Noisy 2D STEM Images
http://openaccess.thecvf.com//content/CVPR2022/html/Kniesel_Clean_Implicit_3D_Structure_From_Noisy_2D_STEM_Images_CVPR_2022_paper.html
AUTHORS: Hannah Kniesel, Timo Ropinski, Tim Bergner, Kavitha Shaga Devan, Clarissa Read, Paul Walther, Tobias Ritschel, Pedro Hermosilla
HIGHLIGHT: Unfortunately, these 2D images can be too noisy to be fused into a useful 3D structure and facilitating good denoisers is challenging due to the lack of clean-noisy pairs. Additionally, representing detailed 3D structure can be difficult even for clean data when using regular 3D grids. Addressing these two limitations, we suggest a differentiable image formation model for STEM, allowing to learn a joint model of 2D sensor noise in STEM together with an implicit 3D model.

1929, TITLE: UKPGAN: A General Self-Supervised Keypoint Detector

http://openaccess.thecvf.com//content/CVPR2022/html/You_UKPGAN_A_General_Self-Supervised_Keypoint_Detector_CVPR_2022_paper.html
AUTHORS: Yang You, Wenhai Liu, Yanjie Ze, Yong-Lu Li, Weiming Wang, Cewu Lu
HIGHLIGHT: In this work, we reckon keypoint detection as information compression, and force the model to distill out important points of an object.

1930, TITLE: Learning Optimal K-Space Acquisition and Reconstruction Using Physics-Informed Neural Networks
http://openaccess.thecvf.com//content/CVPR2022/html/Peng_Learning_Optimal_K-Space_Acquisition_and_Reconstruction_Using_Physics-Informed_Neural_Networks_CVPR_2022_paper.html
AUTHORS: Wei Peng, Li Feng, Guoying Zhao, Fang Liu
HIGHLIGHT: In this work, we propose a novel framework to learn optimized k-space sampling trajectories using deep learning by considering it as an Ordinary Differential Equation (ODE) problem that can be solved using neural ODE.

1931, TITLE: Leveraging Adversarial Examples To Quantify Membership Information Leakage
http://openaccess.thecvf.com//content/CVPR2022/html/Del_Grosso_Leveraging_Adversarial_Examples_To_Quantify_Membership_Information_Leakage_CVPR_2022_paper.html
AUTHORS: Ganesh Del Grosso, Hamid Jalalzai, Georg Pichler, Catuscia Palamidessi, Pablo Piantanida
HIGHLIGHT: The strategy we propose consists of measuring the magnitude of a perturbation necessary to build an adversarial example.

1932, TITLE: Raw High-Definition Radar for Multi-Task Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Rebut_Raw_High-Definition_Radar_for_Multi-Task_Learning_CVPR_2022_paper.html
AUTHORS: Julien Rebut, Arthur Ouaknine, Waqas Malik, Patrick Pérez
HIGHLIGHT: In this paper, we propose a novel HD radar sensing model, FFT-RadNet, that eliminates the overhead of computing the range-azimuth-Doppler 3D tensor, learning instead to recover angles from a range-Doppler spectrum.

1933, TITLE: Point-NeRF: Point-Based Neural Radiance Fields
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_Point-NeRF_Point-Based_Neural_Radiance_Fields_CVPR_2022_paper.html
AUTHORS: Qiangeng Xu, Zexiang Xu, Julien Philip, Sai Bi, Zhixin Shu, Kalyan Sunkavalli, Ulrich Neumann
HIGHLIGHT: Volumetric neural rendering methods like NeRF generate high-quality view synthesis results but are optimized per-scene leading to prohibitive reconstruction time. On the other hand, deep multi-view stereo methods can quickly reconstruct scene geometry via direct network inference. Point-NeRF combines the advantages of these two approaches by using neural 3D point clouds, with associated neural features, to model a radiance field.

1934, TITLE: Contextual Debiasing for Visual Recognition With Causal Mechanisms
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Contextual_Debiasing_for_Visual_Recognition_With_Causal_Mechanisms_CVPR_2022_paper.html
AUTHORS: Ruyang Liu, Hao Liu, Ge Li, Haodi Hou, TingHao Yu, Tao Yang
HIGHLIGHT: In this paper, we propose a simple but effective framework employing causal inference to mitigate contextual bias.

1935, TITLE: Complex Video Action Reasoning via Learnable Markov Logic Network
http://openaccess.thecvf.com//content/CVPR2022/html/Jin_Complex_Video_Action_Reasoning_via_Learnable_Markov_Logic_Network_CVPR_2022_paper.html
AUTHORS: Yang Jin, Linchao Zhu, Yadong Mu
HIGHLIGHT: The main contributions of this work are two-fold: 1) Different from existing black-box models, the proposed model simultaneously implements the localization of temporal boundaries and the recognition of action categories by grounding the logical rules of MLN in videos.

1936, TITLE: Per-Clip Video Object Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Park_Per-Clip_Video_Object_Segmentation_CVPR_2022_paper.html
AUTHORS: Kwanyong Park, Sanghyun Woo, Seoung Wug Oh, In So Kweon, Joon-Young Lee
HIGHLIGHT: The scheme provides two potential benefits: accuracy gain by clip-level optimization and efficiency gain by parallel computation of multiple frames. To this end, we propose a new method tailored for the per-clip inference.

1937, TITLE: Exploring Set Similarity for Dense Self-Supervised Representation Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Exploring_Set_Similarity_for_Dense_Self-Supervised_Representation_Learning_CVPR_2022_paper.html
AUTHORS: Zhaoqing Wang, Qiang Li, Guoxin Zhang, Pengfei Wan, Wen Zheng, Nannan Wang, Mingming Gong, Tongliang Liu
HIGHLIGHT: However, the pixel-level correspondence tends to be noisy because of many similar misleading pixels, e.g., backgrounds. To address this issue, in this paper, we propose to explore set similarity (SetSim) for dense self-supervised representation learning.

1938, TITLE: Coarse-To-Fine Feature Mining for Video Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Sun_Coarse-To-Fine_Feature_Mining_for_Video_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Guolei Sun, Yun Liu, Henghui Ding, Thomas Probst, Luc Van Gool

HIGHLIGHT: However, there is no research about how to simultaneously learn static and motional contexts which are highly correlated and complementary to each other. To address this problem, we propose a Coarse-to-Fine Feature Mining (CFFM) technique to learn a unified presentation of static contexts and motional contexts.

1939, TITLE: ONCE-3DLanes: Building Monocular 3D Lane Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Yan_ONCE-3DLanes_Building_Monocular_3D_Lane_Detection_CVPR_2022_paper.html
AUTHORS: Fan Yan, Ming Nie, Xinyue Cai, Jianhua Han, Hang Xu, Zhen Yang, Chaoqiang Ye, Yanwei Fu, Michael Bi Mi, Li Zhang
HIGHLIGHT: However, existing 3D lane detection datasets are either unpublished or synthesized from a simulated environment, severely hampering the development of this field. In this paper, we take steps towards addressing these issues.

1940, TITLE: Weakly but Deeply Supervised Occlusion-Reasoned Parametric Road Layouts
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Weakly_but_Deeply_Supervised_Occlusion-Reasoned_Parametric_Road_Layouts_CVPR_2022_paper.html
AUTHORS: Buyu Liu, Bingbing Zhuang, Manmohan Chandraker
HIGHLIGHT: We propose an end-to-end network that takes a single perspective RGB image of a complex road scene as input, to produce occlusion-reasoned layouts in perspective space as well as a parametric bird's-eye-view (BEV) space.

1941, TITLE: Compressing Models With Few Samples: Mimicking Then Replacing
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Compressing_Models_With_Few_Samples_Mimicking_Then_Replacing_CVPR_2022_paper.html
AUTHORS: Huanyu Wang, Junjie Liu, Xin Ma, Yang Yong, Zhenhua Chai, Jianxin Wu
HIGHLIGHT: In this paper, we propose a new framework named Mimicking then Replacing (MiR) for few-sample compression, which firstly urges the pruned model to output the same features as the teacher's in the penultimate layer, and then replaces teacher's layers before penultimate with a well-tuned compact one.

1942, TITLE: FedCor: Correlation-Based Active Client Selection Strategy for Heterogeneous Federated Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Tang_FedCor_Correlation-Based_Active_Client_Selection_Strategy_for_Heterogeneous_Federated_Learning_CVPR_2022_paper.html
AUTHORS: Minxue Tang, Xuefei Ning, Yitu Wang, Jingwei Sun, Yu Wang, Hai Li, Yiran Chen
HIGHLIGHT: In this work, we propose FedCor---an FL framework built on a correlation-based client selection strategy, to boost the convergence rate of FL.

1943, TITLE: Modulated Contrast for Versatile Image Synthesis
http://openaccess.thecvf.com//content/CVPR2022/html/Zhan_Modulated_Contrast_for_Versatile_Image_Synthesis_CVPR_2022_paper.html
AUTHORS: Fangneng Zhan, Jiahui Zhang, Yingchen Yu, Rongliang Wu, Shijian Lu
HIGHLIGHT: This paper presents MoNCE, a versatile metric that introduces image contrast to learn a calibrated metric for the perception of multifaceted inter-image distances.

1944, TITLE: PokeBNN: A Binary Pursuit of Lightweight Accuracy
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_PokeBNN_A_Binary_Pursuit_of_Lightweight_Accuracy_CVPR_2022_paper.html
AUTHORS: Yichi Zhang, Zhiru Zhang, Lukasz Lew
HIGHLIGHT: Binary neural networks (BNNs) have the potential to significantly lower the compute intensity but existing models suffer from low quality. To overcome this deficiency, we propose PokeConv, a binary convolution block which improves quality of BNNs by techniques such as adding multiple residual paths, and tuning the activation function.

1945, TITLE: HumanNeRF: Efficiently Generated Human Radiance Field From Sparse Inputs
http://openaccess.thecvf.com//content/CVPR2022/html/Zhao_HumanNeRF_Efficiently_Generated_Human_Radiance_Field_From_Sparse_Inputs_CVPR_2022_paper.html
AUTHORS: Fuqiang Zhao, Wei Yang, Jiakai Zhang, Pei Lin, Yingliang Zhang, Jingyi Yu, Lan Xu
HIGHLIGHT: We present HumanNeRF - a neural representation with efficient generalization ability - for high-fidelity free-view synthesis of dynamic humans.

1946, TITLE: Zoom in and Out: A Mixed-Scale Triplet Network for Camouflaged Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Pang_Zoom_in_and_Out_A_Mixed-Scale_Triplet_Network_for_Camouflaged_CVPR_2022_paper.html
AUTHORS: Youwei Pang, Xiaoqi Zhao, Tian-Zhu Xiang, Lihe Zhang, Huchuan Lu
HIGHLIGHT: Apart from high intrinsic similarity between the camouflaged objects and their background, the objects are usually diverse in scale, fuzzy in appearance, and even severely occluded. To deal with these problems, we propose a mixed-scale triplet network, ZoomNet, which mimics the behavior of humans when observing vague images, i.e., zooming in and out.

1947, TITLE: Identifying Ambiguous Similarity Conditions via Semantic Matching
http://openaccess.thecvf.com//content/CVPR2022/html/Ye_Identifying_Ambiguous_Similarity_Conditions_via_Semantic_Matching_CVPR_2022_paper.html
AUTHORS: Han-Jia Ye, Yi Shi, De-Chuan Zhan
HIGHLIGHT: For example, the previous comparison becomes invalid once the conditional label changes to "is vehicle". To this end, we introduce a novel evaluation criterion by predicting the comparison's correctness after assigning the

learned embeddings to their optimal conditions, which measures how much WS-CSL could cover latent semantics as the supervised model.

1948, TITLE: MISF: Multi-Level Interactive Siamese Filtering for High-Fidelity Image Inpainting
http://openaccess.thecvf.com//content/CVPR2022/html/Li_MISF_Multi-Level_Interactive_Siamese_Filtering_for_High-Fidelity_Image_Inpainting_CVPR_2022_paper.html
AUTHORS: Xiaoguang Li, Qing Guo, Di Lin, Ping Li, Wei Feng, Song Wang
HIGHLIGHT: To address the issues while adopting the respective advantages, we propose a novel filtering technique, i.e., Multi-level Interactive Siamese Filtering (MISF) containing two branches: kernel prediction branch (KPB) and semantic & image filtering branch (SIFB).

1949, TITLE: Cascade Transformers for End-to-End Person Search
http://openaccess.thecvf.com//content/CVPR2022/html/Yu_Cascade_Transformers_for_End-to-End_Person_Search_CVPR_2022_paper.html
AUTHORS: Rui Yu, Dawei Du, Rodney LaLonde, Daniel Davila, Christopher Funk, Anthony Hoogs, Brian Clipp
HIGHLIGHT: In this paper, we propose the Cascade Occluded Attention Transformer (COAT) for end-to-end person search.

1950, TITLE: MSTR: Multi-Scale Transformer for End-to-End Human-Object Interaction Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Kim_MSTR_Multi-Scale_Transformer_for_End-to-End_Human-Object_Interaction_Detection_CVPR_2022_paper.html
AUTHORS: Bumsoo Kim, Jonghwan Mun, Kyoung-Woon On, Minchul Shin, Junhyun Lee, Eun-Sol Kim
HIGHLIGHT: However, they are limited to single-scale feature resolution, providing suboptimal performance in scenes containing humans, objects, and their interactions with vastly different scales and distances. To tackle this problem, we propose a Multi-Scale TRansformer (MSTR) for HOI detection powered by two novel HOI-aware deformable attention modules called Dual-Entity attention and Entity-conditioned Context attention.

1951, TITLE: LSVC: A Learning-Based Stereo Video Compression Framework
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_LSVC_A_Learning-Based_Stereo_Video_Compression_Framework_CVPR_2022_paper.html
AUTHORS: Zhenghao Chen, Guo Lu, Zhihao Hu, Shan Liu, Wei Jiang, Dong Xu
HIGHLIGHT: In this work, we propose the first end-to-end optimized framework for compressing automotive stereo videos (i.e., stereo videos from autonomous driving applications) from both left and right views.

1952, TITLE: How Do You Do It? Fine-Grained Action Understanding With Pseudo-Adverbs
http://openaccess.thecvf.com//content/CVPR2022/html/Doughty_How_Do_You_Do_It_Fine-Grained_Action_Understanding_With_Pseudo-Adverbs_CVPR_2022_paper.html
AUTHORS: Hazel Doughty, Cees G. M. Snoek
HIGHLIGHT: We aim to understand how actions are performed and identify subtle differences, such as 'fold firmly' vs. 'fold gently'. To this end, we propose a method which recognizes adverbs across different actions.

1953, TITLE: InsetGAN for Full-Body Image Generation
http://openaccess.thecvf.com//content/CVPR2022/html/Fruhstuck_InsetGAN_for_Full-Body_Image_Generation_CVPR_2022_paper.html
AUTHORS: Anna Frühstück, Krishna Kumar Singh, Eli Shechtman, Niloy J. Mitra, Peter Wonka, Jingwan Lu
HIGHLIGHT: While GANs can produce photo-realistic images in ideal conditions for certain domains, the generation of full-body human images remains difficult due to the diversity of identities, hairstyles, clothing, and the variance in pose. Instead of modeling this complex domain with a single GAN, we propose a novel method to combine multiple pretrained GANs, where one GAN generates a global canvas (e.g., human body) and a set of specialized GANs, or insets, focus on different parts (e.g., faces, shoes) that can be seamlessly inserted onto the global canvas.

1954, TITLE: DetectorDetective: Investigating the Effects of Adversarial Examples on Object Detectors
http://openaccess.thecvf.com//content/CVPR2022/html/Vellaichamy_DetectorDetective_Investigating_the_Effects_of_Adversarial_Examples_on_Object_Detectors_CVPR_2022_paper.html
AUTHORS: Sivapriya Vellaichamy, Matthew Hull, Zijie J. Wang, Nilaksh Das, ShengYun Peng, Haekyu Park, Duen Horng (Polo) Chau
HIGHLIGHT: We propose DetectorDetective, an interactive visual tool that aims to help users better understand the behaviors of a model as adversarial images journey through an object detector.

1955, TITLE: SOMSI: Spherical Novel View Synthesis With Soft Occlusion Multi-Sphere Images
http://openaccess.thecvf.com//content/CVPR2022/html/Habtegebrial_SOMSI_Spherical_Novel_View_Synthesis_With_Soft_Occlusion_Multi-Sphere_Images_CVPR_2022_paper.html
AUTHORS: Tewodros Habtegebrial, Christiano Gava, Marcel Rogge, Didier Stricker, Varun Jampani
HIGHLIGHT: We propose a novel MSI representation called Soft Occlusion MSI (SOMSI) that enables modelling high-dimensional appearance features in MSI while retaining the fast rendering times of a standard MSI.

1956, TITLE: EMScore: Evaluating Video Captioning via Coarse-Grained and Fine-Grained Embedding Matching
http://openaccess.thecvf.com//content/CVPR2022/html/Shi_EMScore_Evaluating_Video_Captioning_via_Coarse-Grained_and_Fine-Grained_Embedding_Matching_CVPR_2022_paper.html
AUTHORS: Yaya Shi, Xu Yang, Haiyang Xu, Chunfeng Yuan, Bing Li, Weiming Hu, Zheng-Jun Zha

HIGHLIGHT: Inspired by human evaluation, we propose EMScore (Embedding Matching-based score), a novel reference-free metric for video captioning, which directly measures similarity between video and candidate captions.

1957, TITLE: SNR-Aware Low-Light Image Enhancement
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_SNR-Aware_Low-Light_Image_Enhancement_CVPR_2022_paper.html
AUTHORS: Xiaogang Xu, Ruixing Wang, Chi-Wing Fu, Jiaya Jia
HIGHLIGHT: This paper presents a new solution for low-light image enhancement by collectively exploiting Signal-to-Noise-Ratio-aware transformers and convolutional models to dynamically enhance pixels with spatial-varying operations.

1958, TITLE: 3D Common Corruptions and Data Augmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Kar_3D_Common_Corruptions_and_Data_Augmentation_CVPR_2022_paper.html
AUTHORS: OÄŸuzhan Fatih Kar, Teresa Yeo, Andrei Atanov, Amir Zamir
HIGHLIGHT: We introduce a set of image transformations that can be used as corruptions to evaluate the robustness of models as well as data augmentation mechanisms for training neural networks.

1959, TITLE: PoseTriplet: Co-Evolving 3D Human Pose Estimation, Imitation, and Hallucination Under Self-Supervision
http://openaccess.thecvf.com//content/CVPR2022/html/Gong_PoseTriplet_Co-Evolving_3D_Human_Pose_Estimation_Imitation_and_Hallucination_Under_CVPR_2022_paper.html
AUTHORS: Kehong Gong, Bingbing Li, Jianfeng Zhang, Tao Wang, Jing Huang, Michael Bi Mi, Jiashi Feng, Xinchao Wang
HIGHLIGHT: In this paper, we propose a novel self-supervised approach that allows us to explicitly generate 2D-3D pose pairs for augmenting supervision, through a self-enhancing dual-loop learning framework.

1960, TITLE: Injecting Semantic Concepts Into End-to-End Image Captioning
http://openaccess.thecvf.com//content/CVPR2022/html/Fang_Injecting_Semantic_Concepts_Into_End-to-End_Image_Captioning_CVPR_2022_paper.html
AUTHORS: Zhiyuan Fang, Jianfeng Wang, Xiaowei Hu, Lin Liang, Zhe Gan, Lijuan Wang, Yezhou Yang, Zicheng Liu
HIGHLIGHT: In this paper, we are concerned with a better-performing detector-free image captioning model, and propose a pure vision transformer-based image captioning model, dubbed as ViTCAP, in which grid representations are used without extracting the regional features.

1961, TITLE: An Efficient Training Approach for Very Large Scale Face Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_An_Efficient_Training_Approach_for_Very_Large_Scale_Face_Recognition_CVPR_2022_paper.html
AUTHORS: Kai Wang, Shuo Wang, Panpan Zhang, Zhipeng Zhou, Zheng Zhu, Xiaobo Wang, Xiaojiang Peng, Baigui Sun, Hao Li, Yang You
HIGHLIGHT: The heavy computational and memory costs mainly result from the million-level dimensionality of thefully connected (FC) layer. To this end, we propose a novel training approach, termed Faster Face Classification (F2C), to alleviate time and cost without sacrificing the performance.

1962, TITLE: Long-Term Video Frame Interpolation via Feature Propagation
http://openaccess.thecvf.com//content/CVPR2022/html/Argaw_Long-Term_Video_Frame_Interpolation_via_Feature_Propagation_CVPR_2022_paper.html
AUTHORS: Dawit Mureja Argaw, In So Kweon
HIGHLIGHT: Hence, VFI works perform well for small frame gaps and perform poorly as the frame gap increases. In this work, we propose a novel framework to address this problem.

1963, TITLE: Coarse-To-Fine Q-Attention: Efficient Learning for Visual Robotic Manipulation via Discretisation
http://openaccess.thecvf.com//content/CVPR2022/html/James_Coarse-To-Fine_Q-Attention_Efficient_Learning_for_Visual_Robotic_Manipulation_via_Discretisation_CVPR_2022_paper.html
AUTHORS: Stephen James, Kentaro Wada, Tristan Laidlow, Andrew J. Davison
HIGHLIGHT: We present a coarse-to-fine discretisation method that enables the use of discrete reinforcement learning approaches in place of unstable and data-inefficient actor-critic methods in continuous robotics domains.

1964, TITLE: Event-Aided Direct Sparse Odometry
http://openaccess.thecvf.com//content/CVPR2022/html/Hidalgo-Carrio_Event-Aided_Direct_Sparse_Odometry_CVPR_2022_paper.html
AUTHORS: Javier Hidalgo-Carrió, Guillermo Gallego, Davide Scaramuzza
HIGHLIGHT: We introduce EDS, a direct monocular visual odometry using events and frames.

1965, TITLE: Group Contextualization for Video Recognition
http://openaccess.thecvf.com//content/CVPR2022/html/Hao_Group_Contextualization_for_Video_Recognition_CVPR_2022_paper.html
AUTHORS: Yanbin Hao, Hao Zhang, Chong-Wah Ngo, Xiangnan He
HIGHLIGHT: In this paper, we propose an efficient feature refinement method that decomposes the feature channels into several groups and separately refines them with different axial contexts in parallel.

1966, TITLE: Single-Domain Generalized Object Detection in Urban Scene via Cyclic-Disentangled Self-Distillation

http://openaccess.thecvf.com//content/CVPR2022/html/Wu_Single-
Domain_Generalized_Object_Detection_in_Urban_Scene_via_Cyclic-Disentangled_Self-Distillation_CVPR_2022_paper.html
AUTHORS: Aming Wu, Cheng Deng
HIGHLIGHT: In this paper, we are concerned with enhancing the generalization capability of object detectors.

1967, TITLE: Visual Abductive Reasoning
http://openaccess.thecvf.com//content/CVPR2022/html/Liang_Visual_Abductive_Reasoning_CVPR_2022_paper.html
AUTHORS: Chen Liang, Wenguan Wang, Tianfei Zhou, Yi Yang
HIGHLIGHT: In this paper, we propose a new task and dataset, Visual Abductive Reasoning (VAR), for examining abductive
reasoning ability of machine intelligence in everyday visual situations.

1968, TITLE: L2G: A Simple Local-to-Global Knowledge Transfer Framework for Weakly Supervised Semantic
Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Jiang_L2G_A_Simple_Local-to-
Global_Knowledge_Transfer_Framework_for_Weakly_Supervised_CVPR_2022_paper.html
AUTHORS: Peng-Tao Jiang, Yuqi Yang, Qibin Hou, Yunchao Wei
HIGHLIGHT: In this paper, we present L2G, a simple online local-to-global knowledge transfer framework for high-quality
object attention mining.

1969, TITLE: Rethinking Bayesian Deep Learning Methods for Semi-Supervised Volumetric Medical Image Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Rethinking_Bayesian_Deep_Learning_Methods_for_Semi-
Supervised_Volumetric_Medical_Image_CVPR_2022_paper.html
AUTHORS: Jianfeng Wang, Thomas Lukasiewicz
HIGHLIGHT: However, unifying the overall architecture under the Bayesian perspective can make the architecture have a
rigorous theoretical basis, so that each part of the architecture can have a clear probabilistic interpretation. Therefore, to solve the
problems, we propose a new generative Bayesian deep learning (GBDL) architecture.

1970, TITLE: Continual Learning With Lifelong Vision Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Continual_Learning_With_Lifelong_Vision_Transformer_CVPR_2022_paper.html
AUTHORS: Zhen Wang, Liu Liu, Yiqun Duan, Yajing Kong, Dacheng Tao
HIGHLIGHT: In this paper, we propose a novel attention-based framework Lifelong Vision Transformer (LVT), to achieve a
better stability-plasticity trade-off for continual learning.

1971, TITLE: MPViT: Multi-Path Vision Transformer for Dense Prediction
http://openaccess.thecvf.com//content/CVPR2022/html/Lee_MPViT_Multi-
Path_Vision_Transformer_for_Dense_Prediction_CVPR_2022_paper.html
AUTHORS: Youngwan Lee, Jonghee Kim, Jeffrey Willette, Sung Ju Hwang
HIGHLIGHT: In this work, with a different perspective from existing Transformers, we explore multi-scale patch embedding
and multi-path structure, constructing the Multi-Path Vision Transformer (MPViT).

1972, TITLE: NICGSlowDown: Evaluating the Efficiency Robustness of Neural Image Caption Generation Models
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_NICGSlowDown_Evaluating_the_Efficiency_Robustness_of_Neural_Image_Caption_Generation_CVPR_2022_paper.html
AUTHORS: Simin Chen, Zihe Song, Mirazul Haque, Cong Liu, Wei Yang
HIGHLIGHT: To further understand such efficiency-oriented threats, we propose a new attack approach, NICGSlowDown, to
evaluate the efficiency robustness of NICG models.

1973, TITLE: Keypoint Transformer: Solving Joint Identification in Challenging Hands and Object Interactions for Accurate
3D Pose Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Hampali_Keypoint_Transformer_Solving_Joint_Identification_in_Challenging_Hands_and_Object_CVPR_2022_paper.html
AUTHORS: Shreyas Hampali, Sayan Deb Sarkar, Mahdi Rad, Vincent Lepetit
HIGHLIGHT: We propose a robust and accurate method for estimating the 3D poses of two hands in close interaction from a
single color image.

1974, TITLE: SemanticStyleGAN: Learning Compositional Generative Priors for Controllable Image Synthesis and Editing
http://openaccess.thecvf.com//content/CVPR2022/html/Shi_SemanticStyleGAN_Learning_Compositional_Generative_Priors_for_Controllable_Image_Synthesis_and_CVPR_2022_paper.html
AUTHORS: Yichun Shi, Xiao Yang, Yangyue Wan, Xiaohui Shen
HIGHLIGHT: We present SemanticStyleGAN, where a generator is trained to model local semantic parts separately and
synthesizes images in a compositional way.

1975, TITLE: Accurate 3D Body Shape Regression Using Metric and Semantic Attributes
http://openaccess.thecvf.com//content/CVPR2022/html/Choutas_Accurate_3D_Body_Shape_Regression_Using_Metric_and_Semantic_Attributes_CVPR_2022_paper.html
AUTHORS: Vasileios Choutas, Lea Müller, Chun-Hao P. Huang, Siyu Tang, Dimitrios Tzionas, Michael J. Black
HIGHLIGHT: We exploit the anthropometric measurements and linguistic shape attributes in several novel ways to train a
neural network, called SHAPY, that regresses 3D human pose and shape from an RGB image.

1976, TITLE: VL-InterpreT: An Interactive Visualization Tool for Interpreting Vision-Language Transformers
http://openaccess.thecvf.com/content/CVPR2022/html/Aflalo_VL-
InterpreT_An_Interactive_Visualization_Tool_for_Interpreting_Vision-Language_Transformers_CVPR_2022_paper.html
AUTHORS: Estelle Aflalo, Meng Du, Shao-Yen Tseng, Yongfei Liu, Chenfei Wu, Nan Duan, Vasudev Lal
HIGHLIGHT: In this paper, we demonstrate the functionalities of VL-InterpreT through the analysis of KD-VLP, an end-to-end pretraining vision-language multimodal transformer-based model, in the tasks of Visual Commonsense Reasoning (VCR) and WebQA, two visual question answering benchmarks.

1977, TITLE: Label-Only Model Inversion Attacks via Boundary Repulsion
http://openaccess.thecvf.com//content/CVPR2022/html/Kahla_Label-
Only_Model_Inversion_Attacks_via_Boundary_Repulsion_CVPR_2022_paper.html
AUTHORS: Mostafa Kahla, Si Chen, Hoang Anh Just, Ruoxi Jia
HIGHLIGHT: In this paper, we introduce an algorithm, Boundary-Repelling Model Inversion (BREP-MI), to invert private training data using only the target model's predicted labels.

1978, TITLE: Privacy-Preserving Online AutoML for Domain-Specific Face Detection
http://openaccess.thecvf.com/content/CVPR2022/html/Yan_Privacy-Preserving_Online_AutoML_for_Domain-
Specific_Face_Detection_CVPR_2022_paper.html
AUTHORS: Chenqian Yan, Yuge Zhang, Quanlu Zhang, Yaming Yang, Xinyang Jiang, Yuqing Yang, Baoyuan Wang
HIGHLIGHT: We introduce "HyperFD", a new privacy-preserving online AutoML framework for face detection.

1979, TITLE: Self-Augmented Unpaired Image Dehazing via Density and Depth Decomposition
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Self-
Augmented_Unpaired_Image_Dehazing_via_Density_and_Depth_Decomposition_CVPR_2022_paper.html
AUTHORS: Yang Yang, Chaoyue Wang, Risheng Liu, Lin Zhang, Xiaojie Guo, Dacheng Tao
HIGHLIGHT: In this paper, we propose a self-augmented image dehazing framework, termed D^4 (Dehazing via Decomposing transmission map into Density and Depth) for haze generation and removal.

1980, TITLE: Neural 3D Video Synthesis From Multi-View Video
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Neural_3D_Video_Synthesis_From_Multi-
View_Video_CVPR_2022_paper.html
AUTHORS: Tianye Li, Mira Slavcheva, Michael Zollhöfer, Simon Green, Christoph Lassner, Changil Kim, Tanner Schmidt, Steven Lovegrove, Michael Goesele, Richard Newcombe, Zhaoyang Lv
HIGHLIGHT: We propose a novel approach for 3D video synthesis that is able to represent multi-view video recordings of a dynamic real-world scene in a compact, yet expressive representation that enables high-quality view synthesis and motion interpolation.

1981, TITLE: LiDAR Snowfall Simulation for Robust 3D Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Hahner_LiDAR_Snowfall_Simulation_for_Robust_3D_Object_Detection_CVPR_2022_paper.html
AUTHORS: Martin Hahner, Christos Sakaridis, Mario Bijelic, Felix Heide, Fisher Yu, Dengxin Dai, Luc Van Gool
HIGHLIGHT: In this paper, we address the problem of LiDAR-based 3D object detection under snowfall.

1982, TITLE: Learning Where To Learn in Cross-View Self-Supervised Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Huang_Learning_Where_To_Learn_in_Cross-View_Self-
Supervised_Learning_CVPR_2022_paper.html
AUTHORS: Lang Huang, Shan You, Mingkai Zheng, Fei Wang, Chen Qian, Toshihiko Yamasaki
HIGHLIGHT: In this paper, we present a new approach, Learning Where to Learn (LEWEL), to adaptively aggregate spatial information of features, so that the projected embeddings could be exactly aligned and thus guide the feature learning better.

1983, TITLE: SemAffiNet: Semantic-Affine Transformation for Point Cloud Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_SemAffiNet_Semantic-
Affine_Transformation_for_Point_Cloud_Segmentation_CVPR_2022_paper.html
AUTHORS: Ziyi Wang, Yongming Rao, Xumin Yu, Jie Zhou, Jiwen Lu
HIGHLIGHT: Based on this technique, we propose SemAffiNet for point cloud semantic segmentation, which utilizes the attention mechanism in the Transformer module to implicitly and explicitly capture global structural knowledge within local parts for overall comprehension of each category.

1984, TITLE: Sparse Object-Level Supervision for Instance Segmentation With Pixel Embeddings
http://openaccess.thecvf.com//content/CVPR2022/html/Wolny_Sparse_Object-
Level_Supervision_for_Instance_Segmentation_With_Pixel_Embeddings_CVPR_2022_paper.html
AUTHORS: Adrian Wolny, Qin Yu, Constantin Pape, Anna Kreshuk
HIGHLIGHT: We propose to address the dense annotation bottleneck by introducing a proposal-free segmentation approach based on non-spatial embeddings, which exploits the structure of the learned embedding space to extract individual instances in a differentiable way.

1985, TITLE: How Much More Data Do I Need? Estimating Requirements for Downstream Tasks
http://openaccess.thecvf.com//content/CVPR2022/html/Mahmood_How_Much_More_Data_Do_I_Need_Estimating_Requirements_for_CVPR_2022_paper.html

AUTHORS: Rafid Mahmood, James Lucas, David Acuna, Daiqing Li, Jonah Philion, Jose M. Alvarez, Zhiding Yu, Sanja Fidler, Marc T. Law
HIGHLIGHT: In this work, we consider a broad class of computer vision tasks and systematically investigate a family of functions that generalize the power-law function to allow for better estimation of data requirements.

1986, TITLE: Structural and Statistical Texture Knowledge Distillation for Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Ji_Structural_and_Statistical_Texture_Knowledge_Distillation_for_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Deyi Ji, Haoran Wang, Mingyuan Tao, Jianqiang Huang, Xian-Sheng Hua, Hongtao Lu
HIGHLIGHT: In this paper, we are intended to take full advantage of both structural and statistical texture knowledge and propose a novel Structural and Statistical Texture Knowledge Distillation (SSTKD) framework for Semantic Segmentation.

1987, TITLE: Shapley-NAS: Discovering Operation Contribution for Neural Architecture Search
http://openaccess.thecvf.com//content/CVPR2022/html/Xiao_Shapley-NAS_Discovering_Operation_Contribution_for_Neural_Architecture_Search_CVPR_2022_paper.html
AUTHORS: Han Xiao, Ziwei Wang, Zheng Zhu, Jie Zhou, Jiwen Lu
HIGHLIGHT: In this paper, we propose a Shapley value based method to evaluate operation contribution (Shapley-NAS) for neural architecture search.

1988, TITLE: The Implicit Values of a Good Hand Shake: Handheld Multi-Frame Neural Depth Refinement
http://openaccess.thecvf.com//content/CVPR2022/html/Chugunov_The_Implicit_Values_of_a_Good_Hand_Shake_Handheld_Multi-Frame_CVPR_2022_paper.html
AUTHORS: Ilya Chugunov, Yuxuan Zhang, Zhihao Xia, Xuaner Zhang, Jiawen Chen, Felix Heide
HIGHLIGHT: In this work we explore how, from a bundle of these measurements acquired during viewfinding, we can combine dense micro-baseline parallax cues with kilopixel LiDAR depth to distill a high-fidelity depth map.

1989, TITLE: Learning What Not To Segment: A New Perspective on Few-Shot Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Lang_Learning_What_Not_To_Segment_A_New_Perspective_on_Few-Shot_CVPR_2022_paper.html
AUTHORS: Chunbo Lang, Gong Cheng, Binfei Tu, Junwei Han
HIGHLIGHT: This paper proposes a fresh and straightforward insight to alleviate the problem.

1990, TITLE: Blended Diffusion for Text-Driven Editing of Natural Images
http://openaccess.thecvf.com//content/CVPR2022/html/Avrahami_Blended_Diffusion_for_Text-Driven_Editing_of_Natural_Images_CVPR_2022_paper.html
AUTHORS: Omri Avrahami, Dani Lischinski, Ohad Fried
HIGHLIGHT: In this paper, we introduce the first solution for performing local (region-based) edits in generic natural images, based on a natural language description along with an ROI mask.

1991, TITLE: Towards Unsupervised Domain Generalization
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Towards_Unsupervised_Domain_Generalization_CVPR_2022_paper.html
AUTHORS: Xingxuan Zhang, Linjun Zhou, Renzhe Xu, Peng Cui, Zheyan Shen, Haoxin Liu
HIGHLIGHT: Specifically, we study a novel generalization problem called unsupervised domain generalization (UDG), which aims to learn generalizable models with unlabeled data and analyze the effects of pre-training on DG.

1992, TITLE: HyperTransformer: A Textural and Spectral Feature Fusion Transformer for Pansharpening
http://openaccess.thecvf.com//content/CVPR2022/html/Bandara_HyperTransformer_A_Textural_and_Spectral_Feature_Fusion_Transformer_for_Pansharpening_CVPR_2022_paper.html
AUTHORS: Wele Gedara Chaminda Bandara, Vishal M. Patel
HIGHLIGHT: In this paper, we present a novel attention mechanism for pansharpening called HyperTransformer, in which features of LR-HSI and PAN are formulated as queries and keys in a transformer, respectively.

1993, TITLE: Segment-Fusion: Hierarchical Context Fusion for Robust 3D Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Thyagharajan_Segment-Fusion_Hierarchical_Context_Fusion_for_Robust_3D_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Anirud Thyagharajan, Benjamin Ummenhofer, Prashant Laddha, Om Ji Omer, Sreenivas Subramoney
HIGHLIGHT: This paper presents Segment-Fusion, a novel attention-based method for hierarchical fusion of semantic and instance information to address the part misclassifications.

1994, TITLE: Robust Invertible Image Steganography
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_Robust_Invertible_Image_Steganography_CVPR_2022_paper.html
AUTHORS: Youmin Xu, Chong Mou, Yujie Hu, Jingfen Xie, Jian Zhang
HIGHLIGHT: This paper presents a novel flow-based framework for robust invertible image steganography, dubbed as RIIS.

1995, TITLE: Entropy-Based Active Learning for Object Detection With Progressive Diversity Constraint
http://openaccess.thecvf.com//content/CVPR2022/html/Wu_Entropy-Based_Active_Learning_for_Object_Detection_With_Progressive_Diversity_Constraint_CVPR_2022_paper.html
AUTHORS: Jiaxi Wu, Jiaxin Chen, Di Huang

HIGHLIGHT: Active learning for object detection is more challenging and existing efforts on it are relatively rare. In this paper, we propose a novel hybrid approach to address this problem, where the instance-level uncertainty and diversity are jointly considered in a bottom-up manner.

1996, TITLE: BE-STI: Spatial-Temporal Integrated Network for Class-Agnostic Motion Prediction With Bidirectional Enhancement
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_BE-STI_Spatial-Temporal_Integrated_Network_for_Class-Agnostic_Motion_Prediction_With_Bidirectional_CVPR_2022_paper.html
AUTHORS: Yunlong Wang, Hongyu Pan, Jun Zhu, Yu-Huan Wu, Xin Zhan, Kun Jiang, Diange Yang
HIGHLIGHT: In this paper, we propose a novel Spatial-Temporal Integrated network with Bidirectional Enhancement, BE-STI, to improve the temporal motion prediction performance by spatial semantic features, which points out an efficient way to combine semantic segmentation and motion prediction.

1997, TITLE: A Structured Dictionary Perspective on Implicit Neural Representations
http://openaccess.thecvf.com//content/CVPR2022/html/Yuce_A_Structured_Dictionary_Perspective_on_Implicit_Neural_Representations_CVPR_2022_paper.html
AUTHORS: Gizem Yüce, Guillermo Ortiz-Jiménez, Beril Besbinar, Pascal Frossard
HIGHLIGHT: We propose a novel unified perspective to theoretically analyse INRs.

1998, TITLE: Egocentric Deep Multi-Channel Audio-Visual Active Speaker Localization
http://openaccess.thecvf.com//content/CVPR2022/html/Jiang_Egocentric_Deep_Multi-Channel_Audio-Visual_Active_Speaker_Localization_CVPR_2022_paper.html
AUTHORS: Hao Jiang, Calvin Murdock, Vamsi Krishna Ithapu
HIGHLIGHT: We propose a novel end-to-end deep learning approach that is able to give robust voice activity detection and localization results.

1999, TITLE: Vision-Language Pre-Training With Triple Contrastive Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Vision-Language_Pre-Training_With_Triple_Contrastive_Learning_CVPR_2022_paper.html
AUTHORS: Jinyu Yang, Jiali Duan, Son Tran, Yi Xu, Sampath Chanda, Liqun Chen, Belinda Zeng, Trishul Chilimbi, Junzhou Huang
HIGHLIGHT: In this paper, we propose triple contrastive learning (TCL) for vision-language pre-training by leveraging both cross-modal and intra-modal self-supervision.

2000, TITLE: Structure-Aware Flow Generation for Human Body Reshaping
http://openaccess.thecvf.com//content/CVPR2022/html/Ren_Structure-Aware_Flow_Generation_for_Human_Body_Reshaping_CVPR_2022_paper.html
AUTHORS: Jianqiang Ren, Yuan Yao, Biwen Lei, Miaomiao Cui, Xuansong Xie
HIGHLIGHT: Due to the complicated structure and multifarious appearance of human bodies, existing methods either fall back on the 3D domain via body morphable model or resort to keypoint-based image deformation, leading to inefficiency and unsatisfied visual quality. In this paper, we address these limitations by formulating an end-to-end flow generation architecture under the guidance of body structural priors, including skeletons and Part Affinity Fields, and achieve unprecedentedly controllable performance under arbitrary poses and garments.

2001, TITLE: Practical Learned Lossless JPEG Recompression With Multi-Level Cross-Channel Entropy Model in the DCT Domain
http://openaccess.thecvf.com//content/CVPR2022/html/Guo_Practical_Learned_Lossless_JPEG_Recompression_With_Multi-Level_Cross-Channel_Entropy_Model_CVPR_2022_paper.html
AUTHORS: Lina Guo, Xinjie Shi, Dailan He, Yuanyuan Wang, Rui Ma, Hongwei Qin, Yan Wang
HIGHLIGHT: However, most recent progress on image compression mainly focuses on uncompressed images while ignoring trillions of already-existing JPEG images. To compress these JPEG images adequately and restore them back to JPEG format losslessly when needed, we propose a deep learning based JPEG recompression method that operates on DCT domain and propose a Multi-Level Cross-Channel Entropy Model to compress the most informative Y component.

2002, TITLE: Fourier PlenOctrees for Dynamic Radiance Field Rendering in Real-Time
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Fourier_PlenOctrees_for_Dynamic_Radiance_Field_Rendering_in_Real-Time_CVPR_2022_paper.html
AUTHORS: Liao Wang, Jiakai Zhang, Xinhang Liu, Fuqiang Zhao, Yanshun Zhang, Yingliang Zhang, Minye Wu, Jingyi Yu, Lan Xu
HIGHLIGHT: In this paper, we present a novel Fourier PlenOctree (FPO) technique to tackle efficient neural modeling and real-time rendering of dynamic scenes captured under the free-view video (FVV) setting.

2003, TITLE: Learning To Answer Questions in Dynamic Audio-Visual Scenarios
http://openaccess.thecvf.com//content/CVPR2022/html/Li_Learning_To_Answer_Questions_in_Dynamic_Audio-Visual_Scenarios_CVPR_2022_paper.html
AUTHORS: Guangyao Li, Yake Wei, Yapeng Tian, Chenliang Xu, Ji-Rong Wen, Di Hu
HIGHLIGHT: In this paper, we focus on the Audio-Visual Question Answering (AVQA) task, which aims to answer questions regarding different visual objects, sounds, and their associations in videos.

2004, TITLE: Leveraging Equivariant Features for Absolute Pose Regression

213

http://openaccess.thecvf.com//content/CVPR2022/html/Musallam_Leveraging_Equivariant_Features_for_Absolute_Pose_Regression_CVPR_2022_paper.html

AUTHORS: Mohamed Adel Musallam, Vincent Gaudillière, Miguel Ortiz del Castillo, Kassem Al Ismaeil, Djamila Aouada

HIGHLIGHT: In this paper, we demonstrate how a translation and rotation equivariant Convolutional Neural Network directly induces representations of camera motions into the feature space.

2005, TITLE: Synthetic Aperture Imaging With Events and Frames

http://openaccess.thecvf.com//content/CVPR2022/html/Liao_Synthetic_Aperture_Imaging_With_Events_and_Frames_CVPR_2022_paper.html

AUTHORS: Wei Liao, Xiang Zhang, Lei Yu, Shijie Lin, Wen Yang, Ning Qiao

HIGHLIGHT: However, the performance of E-SAI is not consistent under sparse occlusions due to the dramatic decrease of signal events. This paper addresses this problem by leveraging the merits of both events and frames, leading to a fusion-based SAI (EF-SAI) that performs consistently under the different densities of occlusions.

2006, TITLE: CLIP-Event: Connecting Text and Images With Event Structures

http://openaccess.thecvf.com//content/CVPR2022/html/Li_CLIP-Event_Connecting_Text_and_Images_With_Event_Structures_CVPR_2022_paper.html

AUTHORS: Manling Li, Ruochen Xu, Shuohang Wang, Luowei Zhou, Xudong Lin, Chenguang Zhu, Michael Zeng, Heng Ji, Shih-Fu Chang

HIGHLIGHT: In this work, we propose a contrastive learning framework to enforce vision-language pretraining models to comprehend events and associated argument (participant) roles.

2007, TITLE: MonoGround: Detecting Monocular 3D Objects From the Ground

http://openaccess.thecvf.com//content/CVPR2022/html/Qin_MonoGround_Detecting_Monocular_3D_Objects_From_the_Ground_CVPR_2022_paper.html

AUTHORS: Zequn Qin, Xi Li

HIGHLIGHT: Due to the ill-posed 2D to 3D mapping essence from the monocular imaging process, monocular 3D object detection suffers from inaccurate depth estimation and thus has poor 3D detection results. To alleviate this problem, we propose to introduce the ground plane as a prior in the monocular 3d object detection.

2008, TITLE: Deep Visual Geo-Localization Benchmark

http://openaccess.thecvf.com//content/CVPR2022/html/Berton_Deep_Visual_Geo-Localization_Benchmark_CVPR_2022_paper.html

AUTHORS: Gabriele Berton, Riccardo Mereu, Gabriele Trivigno, Carlo Masone, Gabriela Csurka, Torsten Sattler, Barbara Caputo

HIGHLIGHT: In this paper, we propose a new open-source benchmarking framework for Visual Geo-localization (VG) that allows to build, train, and test a wide range of commonly used architectures, with the flexibility to change individual components of a geo-localization pipeline.

2009, TITLE: Scaling Up Vision-Language Pre-Training for Image Captioning

http://openaccess.thecvf.com//content/CVPR2022/html/Hu_Scaling_Up_Vision-Language_Pre-Training_for_Image_Captioning_CVPR_2022_paper.html

AUTHORS: Xiaowei Hu, Zhe Gan, Jianfeng Wang, Zhengyuan Yang, Zicheng Liu, Yumao Lu, Lijuan Wang

HIGHLIGHT: In this paper, we present LEMON, a LargE-scale iMage captiONer, and provide the first empirical study on the scaling behavior of VLP for image captioning.

2010, TITLE: Semiconductor Defect Detection by Hybrid Classical-Quantum Deep Learning

http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Semiconductor_Defect_Detection_by_Hybrid_Classical-Quantum_Deep_Learning_CVPR_2022_paper.html

AUTHORS: Yuan-Fu Yang, Min Sun

HIGHLIGHT: In this paper, we utilize the information processing advantages of quantum computing to promote the defect learning defect review (DLDR).

2011, TITLE: StyleGAN-V: A Continuous Video Generator With the Price, Image Quality and Perks of StyleGAN2

http://openaccess.thecvf.com//content/CVPR2022/html/Skorokhodov_StyleGAN-V_A_Continuous_Video_Generator_With_the_Price_Image_Quality_CVPR_2022_paper.html

AUTHORS: Ivan Skorokhodov, Sergey Tulyakov, Mohamed Elhoseiny

HIGHLIGHT: Videos show continuous events, yet most -- if not all -- video synthesis frameworks treat them discretely in time. In this work, we think of videos of what they should be -- time-continuous signals, and extend the paradigm of neural representations to build a continuous-time video generator.

2012, TITLE: Towards Practical Deployment-Stage Backdoor Attack on Deep Neural Networks

http://openaccess.thecvf.com//content/CVPR2022/html/Qi_Towards_Practical_Deployment-Stage_Backdoor_Attack_on_Deep_Neural_Networks_CVPR_2022_paper.html

AUTHORS: Xiangyu Qi, Tinghao Xie, Ruizhe Pan, Jifeng Zhu, Yong Yang, Kai Bu

HIGHLIGHT: To approach realistic practicality, we propose the first gray-box and physically realizable weights attack algorithm for backdoor injection, namely subnet replacement attack (SRA), which only requires architecture information of the victim model and can support physical triggers in the real world.

2013, TITLE: Scaling Vision Transformers

http://openaccess.thecvf.com//content/CVPR2022/html/Zhai_Scaling_Vision_Transformers_CVPR_2022_paper.html
AUTHORS: Xiaohua Zhai, Alexander Kolesnikov, Neil Houlsby, Lucas Beyer
HIGHLIGHT: While the laws for scaling Transformer language models have been studied, it is unknown how Vision Transformers scale. To address this, we scale ViT models and data, both up and down, and characterize the relationships between error rate, data, and compute.

2014, TITLE: Unsupervised Action Segmentation by Joint Representation Learning and Online Clustering
http://openaccess.thecvf.com//content/CVPR2022/html/Kumar_Unsupervised_Action_Segmentation_by_Joint_Representation_Learning_and_Online_Clustering_CVPR_2022_paper.html
AUTHORS: Sateesh Kumar, Sanjay Haresh, Awais Ahmed, Andrey Konin, M. Zeeshan Zia, Quoc-Huy Tran
HIGHLIGHT: We present a novel approach for unsupervised activity segmentation which uses video frame clustering as a pretext task and simultaneously performs representation learning and online clustering.

2015, TITLE: Pin the Memory: Learning To Generalize Semantic Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Kim_Pin_the_Memory_Learning_To_Generalize_Semantic_Segmentation_CVPR_2022_paper.html
AUTHORS: Jin Kim, Jiyoung Lee, Jungin Park, Dongbo Min, Kwanghoon Sohn
HIGHLIGHT: In this paper, we present a novel memory-guided domain generalization method for semantic segmentation based on meta-learning framework.

2016, TITLE: LISA: Learning Implicit Shape and Appearance of Hands
http://openaccess.thecvf.com//content/CVPR2022/html/Corona_LISA_Learning_Implicit_Shape_and_Appearance_of_Hands_CVPR_2022_paper.html
AUTHORS: Enric Corona, Tomas Hodan, Minh Vo, Francesc Moreno-Noguer, Chris Sweeney, Richard Newcombe, Lingni Ma
HIGHLIGHT: This paper proposes a do-it-all neural model of human hands, named LISA.

2017, TITLE: DiGS: Divergence Guided Shape Implicit Neural Representation for Unoriented Point Clouds
http://openaccess.thecvf.com//content/CVPR2022/html/Ben-Shabat_DiGS_Divergence_Guided_Shape_Implicit_Neural_Representation_for_Unoriented_Point_CVPR_2022_paper.html
AUTHORS: Yizhak Ben-Shabat, Chamin Hewa Koneputugodage, Stephen Gould
HIGHLIGHT: In this paper, we propose a divergence guided shape representation learning approach that does not require normal vectors as input.

2018, TITLE: Iterative Deep Homography Estimation
http://openaccess.thecvf.com//content/CVPR2022/html/Cao_Iterative_Deep_Homography_Estimation_CVPR_2022_paper.html
AUTHORS: Si-Yuan Cao, Jianxin Hu, Zehua Sheng, Hui-Liang Shen
HIGHLIGHT: We propose Iterative Homography Network, namely IHN, a new deep homography estimation architecture.

2019, TITLE: Semi-Supervised Learning of Semantic Correspondence With Pseudo-Labels
http://openaccess.thecvf.com//content/CVPR2022/html/Kim_Semi-Supervised_Learning_of_Semantic_Correspondence_With_Pseudo-Labels_CVPR_2022_paper.html
AUTHORS: Jiwon Kim, Kwangrok Ryoo, Junyoung Seo, Gyuseong Lee, Daehwan Kim, Hansang Cho, Seungryong Kim
HIGHLIGHT: In this paper, we present a simple, but effective solution for semantic correspondence, called SemiMatch, that learns the networks in a semi-supervised manner by supplementing few ground-truth correspondences via utilization of a large amount of confident correspondences as pseudo-labels.

2020, TITLE: Learned Queries for Efficient Local Attention
http://openaccess.thecvf.com//content/CVPR2022/html/Arar_Learned_Queries_for_Efficient_Local_Attention_CVPR_2022_paper.html
AUTHORS: Moab Arar, Ariel Shamir, Amit H. Bermano
HIGHLIGHT: In this paper, we propose a new shift-invariant local attention layer, called query and attend (QnA), that aggregates the input locally in an overlapping manner, much like convolutions.

2021, TITLE: Stereoscopic Universal Perturbations Across Different Architectures and Datasets
http://openaccess.thecvf.com//content/CVPR2022/html/Berger_Stereoscopic_Universal_Perturbations_Across_Different_Architectures_and_Datasets_CVPR_2022_paper.html
AUTHORS: Zachary Berger, Parth Agrawal, Tian Yu Liu, Stefano Soatto, Alex Wong
HIGHLIGHT: We present a method to craft a single set of perturbations that, when added to any stereo image pair in a dataset, can fool a stereo network to significantly alter the perceived scene geometry.

2022, TITLE: Colar: Effective and Efficient Online Action Detection by Consulting Exemplars
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Colar_Effective_and_Efficient_Online_Action_Detection_by_Consulting_Exemplars_CVPR_2022_paper.html
AUTHORS: Le Yang, Junwei Han, Dingwen Zhang
HIGHLIGHT: This paper develops an effective exemplar-consultation mechanism that first measures the similarity between a frame and exemplary frames, and then aggregates exemplary features based on the similarity weights.

2023, TITLE: AutoGPart: Intermediate Supervision Search for Generalizable 3D Part Segmentation

http://openaccess.thecvf.com//content/CVPR2022/html/Liu_AutoGPart_Intermediate_Supervision_Search_for_Generalizable_3D_Part_Segmentation_CVPR_2022_paper.html
AUTHORS: Xueyi Liu, Xiaomeng Xu, Anyi Rao, Chuang Gan, Li Yi
HIGHLIGHT: Others try to use conventional task-agnostic approaches designed for domain generalization problems with no task prior knowledge considered. To solve the above issues, we propose AutoGPart, a generic method enabling training generalizable 3D part segmentation networks with the task prior considered.

2024, TITLE: DeltaCNN: End-to-End CNN Inference of Sparse Frame Differences in Videos
http://openaccess.thecvf.com//content/CVPR2022/html/Parger_DeltaCNN_End-to-End_CNN_Inference_of_Sparse_Frame_Differences_in_Videos_CVPR_2022_paper.html
AUTHORS: Mathias Parger, Chengcheng Tang, Christopher D. Twigg, Cem Keskin, Robert Wang, Markus Steinberger
HIGHLIGHT: With DeltaCNN, we present a sparse convolutional neural network framework that enables sparse frame-by-frame updates to accelerate video inference in practice.

2025, TITLE: HLRTF: Hierarchical Low-Rank Tensor Factorization for Inverse Problems in Multi-Dimensional Imaging
http://openaccess.thecvf.com//content/CVPR2022/html/Luo_HLRTF_Hierarchical_Low-Rank_Tensor_Factorization_for_Inverse_Problems_in_Multi-Dimensional_CVPR_2022_paper.html
AUTHORS: Yisi Luo, Xi-Le Zhao, Deyu Meng, Tai-Xiang Jiang
HIGHLIGHT: To address the vanishing gradient in extreme scenarios, e.g., structural missing pixels, we introduce a parametric total variation regularization to constrain the DNN parameters and the tensor factor parameters with theoretical analysis.

2026, TITLE: Leveraging Self-Supervision for Cross-Domain Crowd Counting
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Leveraging_Self-Supervision_for_Cross-Domain_Crowd_Counting_CVPR_2022_paper.html
AUTHORS: Weizhe Liu, Nikita Durasov, Pascal Fua
HIGHLIGHT: Unfortunately, due to domain shift, the resulting models generalize poorly on real imagery. We remedy this shortcoming by training with both synthetic images, along with their associated labels, and unlabeled real images.

2027, TITLE: MNSRNet: Multimodal Transformer Network for 3D Surface Super-Resolution
http://openaccess.thecvf.com//content/CVPR2022/html/Xie_MNSRNet_Multimodal_Transformer_Network_for_3D_Surface_Super-Resolution_CVPR_2022_paper.html
AUTHORS: Wuyuan Xie, Tengcong Huang, Miaohui Wang
HIGHLIGHT: Due to the unstructured and irregular nature of 3D object data, it is usually difficult to obtain high-quality surface details and geometry textures at a low cost. In this article, we propose an effective multimodal-driven deep neural network to perform 3D surface super-resolution in 2D normal domain, which is simple, accurate, and robust to the above difficulty.

2028, TITLE: Gaussian Process Modeling of Approximate Inference Errors for Variational Autoencoders
http://openaccess.thecvf.com//content/CVPR2022/html/Kim_Gaussian_Process_Modeling_of_Approximate_Inference_Errors_for_Variational_Autoencoders_CVPR_2022_paper.html
AUTHORS: Minyoung Kim
HIGHLIGHT: In this paper, we address the problem in a completely different way by considering a random inference model, where we model the mean and variance functions of the variational posterior as random Gaussian processes (GP).

2029, TITLE: PlaneMVS: 3D Plane Reconstruction From Multi-View Stereo
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_PlaneMVS_3D_Plane_Reconstruction_From_Multi-View_Stereo_CVPR_2022_paper.html
AUTHORS: Jiachen Liu, Pan Ji, Nitin Bansal, Changjiang Cai, Qingan Yan, Xiaolei Huang, Yi Xu
HIGHLIGHT: We present a novel framework named PlaneMVS for 3D plane reconstruction from multiple input views with known camera poses.

2030, TITLE: Scene Graph Expansion for Semantics-Guided Image Outpainting
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_Scene_Graph_Expansion_for_Semantics-Guided_Image_Outpainting_CVPR_2022_paper.html
AUTHORS: Chiao-An Yang, Cheng-Yo Tan, Wan-Cyuan Fan, Cheng-Fu Yang, Meng-Lin Wu, Yu-Chiang Frank Wang
HIGHLIGHT: In this paper, we address the task of semantics-guided image outpainting, which is to complete an image by generating semantically practical content.

2031, TITLE: SoftGroup for 3D Instance Segmentation on Point Clouds
http://openaccess.thecvf.com//content/CVPR2022/html/Vu_SoftGroup_for_3D_Instance_Segmentation_on_Point_Clouds_CVPR_2022_paper.html
AUTHORS: Thang Vu, Kookhoi Kim, Tung M. Luu, Thanh Nguyen, Chang D. Yoo
HIGHLIGHT: To address the aforementioned problems, this paper proposes a 3D instance segmentation method referred to as SoftGroup by performing bottom-up soft grouping followed by top-down refinement.

2032, TITLE: SharpContour: A Contour-Based Boundary Refinement Approach for Efficient and Accurate Instance Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Zhu_SharpContour_A_Contour-Based_Boundary_Refinement_Approach_for_Efficient_and_Accurate_CVPR_2022_paper.html
AUTHORS: Chenming Zhu, Xuanye Zhang, Yanran Li, Liangdong Qiu, Kai Han, Xiaoguang Han

216

HIGHLIGHT: In this paper, we propose an efficient contour-based boundary refinement approach, named SharpContour, to tackle the segmentation of boundary area.

2033, TITLE: MVS2D: Efficient Multi-View Stereo via Attention-Driven 2D Convolutions
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_MVS2D_Efficient_Multi-View_Stereo_via_Attention-Driven_2D_Convolutions_CVPR_2022_paper.html
AUTHORS: Zhenpei Yang, Zhile Ren, Qi Shan, Qixing Huang
HIGHLIGHT: We present MVS2D, a highly efficient multi-view stereo algorithm that seamlessly integrates multi-view constraints into single-view networks via an attention mechanism.

2034, TITLE: FIBA: Frequency-Injection Based Backdoor Attack in Medical Image Analysis
http://openaccess.thecvf.com//content/CVPR2022/html/Feng_FIBA_Frequency-Injection_Based_Backdoor_Attack_in_Medical_Image_Analysis_CVPR_2022_paper.html
AUTHORS: Yu Feng, Benteng Ma, Jing Zhang, Shanshan Zhao, Yong Xia, Dacheng Tao
HIGHLIGHT: Most existing BA methods are designed to attack natural image classification models, which apply spatial triggers to training images and inevitably corrupt the semantics of poisoned pixels, leading to the failures of attacking dense prediction models. To address this issue, we propose a novel Frequency-Injection based Backdoor Attack method (FIBA) that is capable of delivering attacks in various MIA tasks.

2035, TITLE: Beyond Semantic to Instance Segmentation: Weakly-Supervised Instance Segmentation via Semantic Knowledge Transfer and Self-Refinement
http://openaccess.thecvf.com//content/CVPR2022/html/Kim_Beyond_Semantic_to_Instance_Segmentation_Weakly-Supervised_Instance_Segmentation_via_Semantic_CVPR_2022_paper.html
AUTHORS: Beomyoung Kim, YoungJoon Yoo, Chae Eun Rhee, Junmo Kim
HIGHLIGHT: In this paper, we propose a novel approach including two innovative components.

2036, TITLE: Bridged Transformer for Vision and Point Cloud 3D Object Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Bridged_Transformer_for_Vision_and_Point_Cloud_3D_Object_Detection_CVPR_2022_paper.html
AUTHORS: Yikai Wang, TengQi Ye, Lele Cao, Wenbing Huang, Fuchun Sun, Fengxiang He, Dacheng Tao
HIGHLIGHT: To that end, we propose Bridged Transformer (BrT), an end-to-end architecture for 3D object detection.

2037, TITLE: Deep Constrained Least Squares for Blind Image Super-Resolution
http://openaccess.thecvf.com//content/CVPR2022/html/Luo_Deep_Constrained_Least_Squares_for_Blind_Image_Super-Resolution_CVPR_2022_paper.html
AUTHORS: Ziwei Luo, Haibin Huang, Lei Yu, Youwei Li, Haoqiang Fan, Shuaicheng Liu
HIGHLIGHT: In this paper, we tackle the problem of blind image super-resolution(SR) with a reformulated degradation model and two novel modules.

2038, TITLE: EDTER: Edge Detection With Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Pu_EDTER_Edge_Detection_With_Transformer_CVPR_2022_paper.html
AUTHORS: Mengyang Pu, Yaping Huang, Yuming Liu, Qingji Guan, Haibin Ling
HIGHLIGHT: Recently, vision transformer has shown excellent capability in capturing long-range dependencies. Inspired by this, we propose a novel transformer-based edge detector, Edge Detection TransformER (EDTER), to extract clear and crisp object boundaries and meaningful edges by exploiting the full image context information and detailed local cues simultaneously.

2039, TITLE: Fine-Tuning Global Model via Data-Free Knowledge Distillation for Non-IID Federated Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Fine-Tuning_Global_Model_via_Data-Free_Knowledge_Distillation_for_Non-IID_Federated_CVPR_2022_paper.html
AUTHORS: Lin Zhang, Li Shen, Liang Ding, Dacheng Tao, Ling-Yu Duan
HIGHLIGHT: Instead, we propose a data-free knowledge distillation method to fine-tune the global model in the server (FedFTG), which relieves the issue of direct model aggregation.

2040, TITLE: JIFF: Jointly-Aligned Implicit Face Function for High Quality Single View Clothed Human Reconstruction
http://openaccess.thecvf.com//content/CVPR2022/html/Cao_JIFF_Jointly-Aligned_Implicit_Face_Function_for_High_Quality_Single_View_CVPR_2022_paper.html
AUTHORS: Yukang Cao, Guanying Chen, Kai Han, Wenqi Yang, Kwan-Yee K. Wong
HIGHLIGHT: In this paper, we focus on improving the quality of face in the reconstruction and propose a novel Jointly-aligned Implicit Face Function (JIFF) that combines the merits of the implicit function based approach and model based approach.

2041, TITLE: Deep 3D-to-2D Watermarking: Embedding Messages in 3D Meshes and Extracting Them From 2D Renderings
http://openaccess.thecvf.com//content/CVPR2022/html/Yoo_Deep_3D-to-2D_Watermarking_Embedding_Messages_in_3D_Meshes_and_Extracting_CVPR_2022_paper.html
AUTHORS: Innfarn Yoo, Huiwen Chang, Xiyang Luo, Ondrej Stava, Ce Liu, Peyman Milanfar, Feng Yang
HIGHLIGHT: Unfortunately, retrieving messages from 2D renderings of 3D meshes is still challenging and underexplored. We introduce a novel end-to-end learning framework to solve this problem through: 1) an encoder to covertly embed messages in both mesh geometry and textures; 2) a differentiable renderer to render watermarked 3D objects from different camera angles and under varied lighting conditions; 3) a decoder to recover the messages from 2D rendered images.

2042, TITLE: Beyond a Pre-Trained Object Detector: Cross-Modal Textual and Visual Context for Image Captioning

http://openaccess.thecvf.com//content/CVPR2022/html/Kuo_Beyond_a_Pre-Trained_Object_Detector_Cross-Modal_Textual_and_Visual_Context_CVPR_2022_paper.html
AUTHORS: Chia-Wen Kuo, Zsolt Kira
HIGHLIGHT: The assumption that such outputs can represent all necessary information is unrealistic, especially when the detector is transferred across datasets. In this work, we reason about the graphical model induced by this assumption, and propose to add an auxiliary input to represent missing information such as object relationships.

2043, TITLE: Symmetry-Aware Neural Architecture for Embodied Visual Exploration
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_Symmetry-Aware_Neural_Architecture_for_Embodied_Visual_Exploration_CVPR_2022_paper.html
AUTHORS: Shuang Liu, Takayuki Okatani
HIGHLIGHT: In this paper, we attempt to improve the generalization ability by utilizing the inductive biases available for the task.

2044, TITLE: AirObject: A Temporally Evolving Graph Embedding for Object Identification
http://openaccess.thecvf.com//content/CVPR2022/html/Keetha_AirObject_A_Temporally_Evolving_Graph_Embedding_for_Object_Identification_CVPR_2022_paper.html
AUTHORS: Nikhil Varma Keetha, Chen Wang, Yuheng Qiu, Kuan Xu, Sebastian Scherer
HIGHLIGHT: Furthermore, given the vast distribution of unknown novel objects in the real world, the object identification process must be class-agnostic. In this context, we propose a novel temporal 3D object encoding approach, dubbed AirObject, to obtain global keypoint graph-based embeddings of objects.

2045, TITLE: From Representation to Reasoning: Towards Both Evidence and Commonsense Reasoning for Video Question-Answering
http://openaccess.thecvf.com//content/CVPR2022/html/Li_From_Representation_to_Reasoning_Towards_Both_Evidence_and_Commonsense_Reasoning_CVPR_2022_paper.html
AUTHORS: Jiangtong Li, Li Niu, Liqing Zhang
HIGHLIGHT: To facilitate deeper video understanding towards video reasoning, we present the task of Causal-VidQA, which includes four types of questions ranging from scene description (description) to evidence reasoning (explanation) and commonsense reasoning (prediction and counterfactual).

2046, TITLE: Semantic-Aware Domain Generalized Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Peng_Semantic-Aware_Domain_Generalized_Segmentation_CVPR_2022_paper.html
AUTHORS: Duo Peng, Yinjie Lei, Munawar Hayat, Yulan Guo, Wen Li
HIGHLIGHT: In this paper, we address domain generalized semantic segmentation, where a segmentation model is trained to be domain-invariant without using any target domain data.

2047, TITLE: TransVPR: Transformer-Based Place Recognition With Multi-Level Attention Aggregation
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_TransVPR_Transformer-Based_Place_Recognition_With_Multi-Level_Attention_Aggregation_CVPR_2022_paper.html
AUTHORS: Ruotong Wang, Yanqing Shen, Weiliang Zuo, Sanping Zhou, Nanning Zheng
HIGHLIGHT: In this paper, we introduce a novel holistic place recognition model, TransVPR, based on vision Transformers.

2048, TITLE: DanceTrack: Multi-Object Tracking in Uniform Appearance and Diverse Motion
http://openaccess.thecvf.com//content/CVPR2022/html/Sun_DanceTrack_Multi-Object_Tracking_in_Uniform_Appearance_and_Diverse_Motion_CVPR_2022_paper.html
AUTHORS: Peize Sun, Jinkun Cao, Yi Jiang, Zehuan Yuan, Song Bai, Kris Kitani, Ping Luo
HIGHLIGHT: In response to such bias, we would like to re-emphasize that methods for multi-object tracking should also work when object appearance is not sufficiently discriminative. To this end, we propose a large-scale dataset for multi-human tracking, where humans have similar appearance, diverse motion and extreme articulation.

2049, TITLE: Unsupervised Learning of Debiased Representations With Pseudo-Attributes
http://openaccess.thecvf.com//content/CVPR2022/html/Seo_Unsupervised_Learning_of_Debiased_Representations_With_Pseudo-Attributes_CVPR_2022_paper.html
AUTHORS: Seonguk Seo, Joon-Young Lee, Bohyung Han
HIGHLIGHT: Although existing works often handle this issue using human supervision, the availability of the proper annotations is impractical and even unrealistic. To better tackle this challenge, we propose a simple but effective debiasing technique in an unsupervised manner.

2050, TITLE: Protecting Celebrities From DeepFake With Identity Consistency Transformer
http://openaccess.thecvf.com//content/CVPR2022/html/Dong_Protecting_Celebrities_From_DeepFake_With_Identity_Consistency_Transformer_CVPR_2022_paper.html
AUTHORS: Xiaoyi Dong, Jianmin Bao, Dongdong Chen, Ting Zhang, Weiming Zhang, Nenghai Yu, Dong Chen, Fang Wen, Baining Guo
HIGHLIGHT: In this work we propose Identity Consistency Transformer, a novel face forgery detection method that focuses on high-level semantics, specifically identity information, and detecting a suspect face by finding identity inconsistency in inner and outer face regions.

2051, TITLE: Give Me Your Attention: Dot-Product Attention Considered Harmful for Adversarial Patch Robustness

http://openaccess.thecvf.com//content/CVPR2022/html/Lovisotto_Give_Me_Your_Attention_Dot-
Product_Attention_Considered_Harmful_for_Adversarial_CVPR_2022_paper.html
AUTHORS: Giulio Lovisotto, Nicole Finnie, Mauricio Munoz, Chaithanya Kumar Mummadi, Jan Hendrik Metzen
HIGHLIGHT: In this paper, we show how the global reasoning of (scaled) dot-product attention can be the source of a major vulnerability when confronted with adversarial patch attacks.

2052, TITLE: TubeDETR: Spatio-Temporal Video Grounding With Transformers
http://openaccess.thecvf.com//content/CVPR2022/html/Yang_TubeDETR_Spatio-
Temporal_Video_Grounding_With_Transformers_CVPR_2022_paper.html
AUTHORS: Antoine Yang, Antoine Miech, Josef Sivic, Ivan Laptev, Cordelia Schmid
HIGHLIGHT: We consider the problem of localizing a spatio-temporal tube in a video corresponding to a given text query.

2053, TITLE: KG-SP: Knowledge Guided Simple Primitives for Open World Compositional Zero-Shot Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Karthik_KG-
SP_Knowledge_Guided_Simple_Primitives_for_Open_World_Compositional_Zero-Shot_CVPR_2022_paper.html
AUTHORS: Shyamgopal Karthik, Massimiliano Mancini, Zeynep Akata
HIGHLIGHT: While previous works tackle the problem by learning embeddings for the compositions jointly,here we revisit a simple CZSL baseline and predict the primitives, i.e. states and objects, independently.

2054, TITLE: SLIC: Self-Supervised Learning With Iterative Clustering for Human Action Videos
http://openaccess.thecvf.com//content/CVPR2022/html/Khorasgani_SLIC_Self-
Supervised_Learning_With_Iterative_Clustering_for_Human_Action_Videos_CVPR_2022_paper.html
AUTHORS: Salar Hosseini Khorasgani, Yuxuan Chen, Florian Shkurti
HIGHLIGHT: A typical assumption is that similar clips only occur temporally close within a single video, leading to insufficient examples of motion similarity. To mitigate this, we propose SLIC, a clustering-based self-supervised contrastive learning method for human action videos.

2055, TITLE: CD2-pFed: Cyclic Distillation-Guided Channel Decoupling for Model Personalization in Federated Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Shen_CD2-pFed_Cyclic_Distillation-
Guided_Channel_Decoupling_for_Model_Personalization_in_Federated_CVPR_2022_paper.html
AUTHORS: Yiqing Shen, Yuyin Zhou, Lequan Yu
HIGHLIGHT: In this paper, we propose CD^2-pFed, a novel Cyclic Distillation-guided Channel Decoupling framework, to personalize the global model in FL, under various settings of data heterogeneity.

2056, TITLE: UBnormal: New Benchmark for Supervised Open-Set Video Anomaly Detection
http://openaccess.thecvf.com//content/CVPR2022/html/Acsintoae_UBnormal_New_Benchmark_for_Supervised_Open-
Set_Video_Anomaly_Detection_CVPR_2022_paper.html
AUTHORS: Andra Acsintoae, Andrei Florescu, Mariana-Iuliana Georgescu, Tudor Mare, Paul Sumedrea, Radu Tudor Ionescu, Fahad Shahbaz Khan, Mubarak Shah
HIGHLIGHT: This is a closed-set scenario that fails to test the capability of systems at detecting new anomaly types. To this end, we propose UBnormal, a new supervised open-set benchmark composed of multiple virtual scenes for video anomaly detection.

2057, TITLE: Beyond Cross-View Image Retrieval: Highly Accurate Vehicle Localization Using Satellite Image
http://openaccess.thecvf.com//content/CVPR2022/html/Shi_Beyond_Cross-
View_Image_Retrieval_Highly_Accurate_Vehicle_Localization_Using_Satellite_CVPR_2022_paper.html
AUTHORS: Yujiao Shi, Hongdong Li
HIGHLIGHT: Departing from the conventional wisdom of image retrieval, this paper presents a novel solution that can achieve highly-accurate localization.

2058, TITLE: Closing the Generalization Gap of Cross-Silo Federated Medical Image Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Xu_Closing_the_Generalization_Gap_of_Cross-
Silo_Federated_Medical_Image_Segmentation_CVPR_2022_paper.html
AUTHORS: An Xu, Wenqi Li, Pengfei Guo, Dong Yang, Holger R. Roth, Ali Hatamizadeh, Can Zhao, Daguang Xu, Heng Huang, Ziyue Xu
HIGHLIGHT: In this work, we propose a novel training framework FedSM to avoid the client drift issue and successfully close the generalization gap compared with the centralized training for medical image segmentation tasks for the first time.

2059, TITLE: AKB-48: A Real-World Articulated Object Knowledge Base
http://openaccess.thecvf.com//content/CVPR2022/html/Liu_AKB-48_A_Real-
World_Articulated_Object_Knowledge_Base_CVPR_2022_paper.html
AUTHORS: Liu Liu, Wenqiang Xu, Haoyuan Fu, Sucheng Qian, Qiaojun Yu, Yang Han, Cewu Lu
HIGHLIGHT: To build the AKB-48, we present a fast articulation knowledge modeling (FArM) pipeline, which can fulfill the ArtiKG for an articulated object within 10-15 minutes, and largely reduce the cost for object modeling in the real world.

2060, TITLE: Style-ERD: Responsive and Coherent Online Motion Style Transfer
http://openaccess.thecvf.com//content/CVPR2022/html/Tao_Style-
ERD_Responsive_and_Coherent_Online_Motion_Style_Transfer_CVPR_2022_paper.html
AUTHORS: Tianxin Tao, Xiaohang Zhan, Zhongquan Chen, Michiel van de Panne

HIGHLIGHT: However, for online animation applications, such as real-time avatar animation from motion capture, motions need to be processed as a stream with minimal latency. In this work, we realize a flexible, high-quality motion style transfer method for this setting.

2061, TITLE: Leverage Your Local and Global Representations: A New Self-Supervised Learning Strategy
http://openaccess.thecvf.com//content/CVPR2022/html/Zhang_Leverage_Your_Local_and_Global_Representations_A_New_Self-Supervised_Learning_CVPR_2022_paper.html
AUTHORS: Tong Zhang, Congpei Qiu, Wei Ke, Sabine Süsstrunk, Mathieu Salzmann
HIGHLIGHT: In essence, this strategy ignores the fact that two crops may truly contain different image information, e.g., background and small objects, and thus tends to restrain the diversity of the learned representations. In this work, we address this issue by introducing a new self-supervised learning strategy, LoGo, that explicitly reasons about Lo cal and G l o bal crops.

2062, TITLE: Stratified Transformer for 3D Point Cloud Segmentation
http://openaccess.thecvf.com//content/CVPR2022/html/Lai_Stratified_Transformer_for_3D_Point_Cloud_Segmentation_CVPR_2022_paper.html
AUTHORS: Xin Lai, Jianhui Liu, Li Jiang, Liwei Wang, Hengshuang Zhao, Shu Liu, Xiaojuan Qi, Jiaya Jia
HIGHLIGHT: In this paper, we propose Stratified Transformer that is able to capture long-range contexts and demonstrates strong generalization ability and high performance.

2063, TITLE: NeRF in the Dark: High Dynamic Range View Synthesis From Noisy Raw Images
http://openaccess.thecvf.com//content/CVPR2022/html/Mildenhall_NeRF_in_the_Dark_High_Dynamic_Range_View_Synthesis_From_CVPR_2022_paper.html
AUTHORS: Ben Mildenhall, Peter Hedman, Ricardo Martin-Brualla, Pratul P. Srinivasan, Jonathan T. Barron
HIGHLIGHT: We modify NeRF to instead train directly on linear raw images, preserving the scene's full dynamic range.

2064, TITLE: DArch: Dental Arch Prior-Assisted 3D Tooth Instance Segmentation With Weak Annotations
http://openaccess.thecvf.com//content/CVPR2022/html/Qiu_DArch_Dental_Arch_Prior-Assisted_3D_Tooth_Instance_Segmentation_With_Weak_CVPR_2022_paper.html
AUTHORS: Liangdong Qiu, Chongjie Ye, Pei Chen, Yunbi Liu, Xiaoguang Han, Shuguang Cui
HIGHLIGHT: Existing learning-based methods rely heavily on expensive point-wise annotations. To alleviate this problem, we are the first to explore a low-cost annotation way for 3D tooth instance segmentation, i.e., labeling all tooth centroids and only a few teeth for each dental model.

2065, TITLE: Task Decoupled Framework for Reference-Based Super-Resolution
http://openaccess.thecvf.com//content/CVPR2022/html/Huang_Task_Decoupled_Framework_for_Reference-Based_Super-Resolution_CVPR_2022_paper.html
AUTHORS: Yixuan Huang, Xiaoyun Zhang, Yu Fu, Siheng Chen, Ya Zhang, Yan-Feng Wang, Dazhi He
HIGHLIGHT: Those methods conduct the super-resolution task of the input low-resolution(LR) image and the texture transfer task from the reference image together in one module, easily introducing the interference between LR and reference features. Inspired by this finding, we propose a novel framework, which decouples the two tasks of RefSR, eliminating the interference between the LR image and the reference image.

2066, TITLE: Aug-NeRF: Training Stronger Neural Radiance Fields With Triple-Level Physically-Grounded Augmentations
http://openaccess.thecvf.com//content/CVPR2022/html/Chen_Aug-NeRF_Training_Stronger_Neural_Radiance_Fields_With_Triple-Level_Physically-Grounded_Augmentations_CVPR_2022_paper.html
AUTHORS: Tianlong Chen, Peihao Wang, Zhiwen Fan, Zhangyang Wang
HIGHLIGHT: Inspired by that, we propose Augmented NeRF (Aug-NeRF), which for the first time brings the power of robust data augmentations into regularizing the NeRF training.

2067, TITLE: RGB-Multispectral Matching: Dataset, Learning Methodology, Evaluation
http://openaccess.thecvf.com//content/CVPR2022/html/Tosi_RGB-Multispectral_Matching_Dataset_Learning_Methodology_Evaluation_CVPR_2022_paper.html
AUTHORS: Fabio Tosi, Pierluigi Zama Ramirez, Matteo Poggi, Samuele Salti, Stefano Mattoccia, Luigi Di Stefano
HIGHLIGHT: To tackle the task, we propose a deep learning architecture trained in a self-supervised manner by exploiting a further RGB camera, required only during training data acquisition.

2068, TITLE: Id-Free Person Similarity Learning
http://openaccess.thecvf.com//content/CVPR2022/html/Shuai_Id-Free_Person_Similarity_Learning_CVPR_2022_paper.html
AUTHORS: Bing Shuai, Xinyu Li, Kaustav Kundu, Joseph Tighe
HIGHLIGHT: To this end, we present a contrastive learning framework to learn person similarity without using manually labeled identity annotations.

2069, TITLE: Temporal Complementarity-Guided Reinforcement Learning for Image-to-Video Person Re-Identification
http://openaccess.thecvf.com//content/CVPR2022/html/Wu_Temporal_Complementarity-Guided_Reinforcement_Learning_for_Image-to-Video_Person_Re-Identification_CVPR_2022_paper.html
AUTHORS: Wei Wu, Jiawei Liu, Kecheng Zheng, Qibin Sun, Zheng-Jun Zha
HIGHLIGHT: Existing methods treat it as a cross-modality retrieval task and learn the common latent embeddings from image and video modalities, which are both less effective and efficient due to large modality gap and redundant feature learning by utilizing all video frames. In this work, we first regard this task as point-to-set matching problem identical to human decision process, and

propose a novel Temporal Complementarity-Guided Reinforcement Learning (TCRL) approach for image-to-video person re-identification.

2070, TITLE: Globetrotter: Connecting Languages by Connecting Images
http://openaccess.thecvf.com//content/CVPR2022/html/Suris_Globetrotter_Connecting_Languages_by_Connecting_Images_CVPR_2022_paper.html
AUTHORS: Dídac Surís, Dave Epstein, Carl Vondrick
HIGHLIGHT: We introduce a method that uses visual observations to bridge the gap between languages, rather than relying on parallel corpora or topological properties of the representations.

2071, TITLE: Fairness-Aware Adversarial Perturbation Towards Bias Mitigation for Deployed Deep Models
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Fairness-Aware_Adversarial_Perturbation_Towards_Bias_Mitigation_for_Deployed_Deep_Models_CVPR_2022_paper.html
AUTHORS: Zhibo Wang, Xiaowei Dong, Henry Xue, Zhifei Zhang, Weifeng Chiu, Tao Wei, Kui Ren
HIGHLIGHT: By contrast, we propose a more flexible approach, i.e., fairness-aware adversarial perturbation (FAAP), which learns to perturb input data to blind deployed models on fairness-related features, e.g., gender and ethnicity.

2072, TITLE: Stochastic Backpropagation: A Memory Efficient Strategy for Training Video Models
http://openaccess.thecvf.com//content/CVPR2022/html/Cheng_Stochastic_Backpropagation_A_Memory_Efficient_Strategy_for_Training_Video_Models_CVPR_2022_paper.html
AUTHORS: Feng Cheng, Mingze Xu, Yuanjun Xiong, Hao Chen, Xinyu Li, Wei Li, Wei Xia
HIGHLIGHT: We propose a memory efficient method, named Stochastic Backpropagation (SBP), for training deep neural networks on videos.

2073, TITLE: Semantic-Shape Adaptive Feature Modulation for Semantic Image Synthesis
http://openaccess.thecvf.com//content/CVPR2022/html/Lv_Semantic-Shape_Adaptive_Feature_Modulation_for_Semantic_Image_Synthesis_CVPR_2022_paper.html
AUTHORS: Zhengyao Lv, Xiaoming Li, Zhenxing Niu, Bing Cao, Wangmeng Zuo
HIGHLIGHT: In order to exploit the part-level layouts, we propose a Shape-aware Position Descriptor (SPD) to describe each pixel's positional feature, where object shape is explicitly encoded into the SPD feature.

2074, TITLE: Egocentric Scene Understanding via Multimodal Spatial Rectifier
http://openaccess.thecvf.com//content/CVPR2022/html/Do_Egocentric_Scene_Understanding_via_Multimodal_Spatial_Rectifier_CVPR_2022_paper.html
AUTHORS: Tien Do, Khiem Vuong, Hyun Soo Park
HIGHLIGHT: In this paper, we study a problem of egocentric scene understanding, i.e., predicting depths and surface normals from an egocentric image.

2075, TITLE: Semi-Supervised Semantic Segmentation Using Unreliable Pseudo-Labels
http://openaccess.thecvf.com//content/CVPR2022/html/Wang_Semi-Supervised_Semantic_Segmentation_Using_Unreliable_Pseudo-Labels_CVPR_2022_paper.html
AUTHORS: Yuchao Wang, Haochen Wang, Yujun Shen, Jingjing Fei, Wei Li, Guoqiang Jin, Liwei Wu, Rui Zhao, Xinyi Le
HIGHLIGHT: Hence, such a pixel can be convincingly treated as a negative sample to those most unlikely categories. Based on this insight, we develop an effective pipeline to make sufficient use of unlabeled data.

2076, TITLE: Day-to-Night Image Synthesis for Training Nighttime Neural ISPs
http://openaccess.thecvf.com//content/CVPR2022/html/Punnappurath_Day-to-Night_Image_Synthesis_for_Training_Nighttime_Neural_ISPs_CVPR_2022_paper.html
AUTHORS: Abhijith Punnappurath, Abdullah Abuolaim, Abdelrahman Abdelhamed, Alex Levinshtein, Michael S. Brown
HIGHLIGHT: To address this problem, we propose a method that synthesizes nighttime images from daytime images.

2077, TITLE: Commonality in Natural Images Rescues GANs: Pretraining GANs With Generic and Privacy-Free Synthetic Data
http://openaccess.thecvf.com//content/CVPR2022/html/Baek_Commonality_in_Natural_Images_Rescues_GANs_Pretraining_GANs_With_Generic_CVPR_2022_paper.html
AUTHORS: Kyungjune Baek, Hyunjung Shim
HIGHLIGHT: However, existing studies show that the pretrained model using a single benchmark dataset is not generalized to various target datasets. More importantly, the pretrained model can be vulnerable to copyright or privacy risks as membership inference attack advances. To resolve both issues, we propose an effective and unbiased data synthesizer, namely Primitives-PS, inspired by the generic characteristics of natural images.