

- 1, TITLE: CATER: A diagnostic dataset for Compositional Actions & Temporal Reasoning  
<https://openreview.net/forum?id=HJgz2VKPB>  
AUTHORS: Rohit Girdhar, Deva Ramanan  
HIGHLIGHT: We propose a new video understanding benchmark, with tasks that by-design require temporal reasoning to be solved, unlike most existing video datasets.
- 2, TITLE: BackPACK: Packing more into Backprop  
<https://openreview.net/forum?id=BJlrF24twB>  
AUTHORS: Felix Dangel, Frederik Kunstner, Philipp Hennig  
HIGHLIGHT: We propose a new video understanding benchmark, with tasks that by-design require temporal reasoning to be solved, unlike most existing video datasets.
- 3, TITLE: GenDICE: Generalized Offline Estimation of Stationary Values  
<https://openreview.net/forum?id=HkxlcnVFwB>  
AUTHORS: Ruiyi Zhang\*, Bo Dai\*, Lihong Li, Dale Schuurmans  
HIGHLIGHT: In this paper, we proposed a novel algorithm, GenDICE, for general stationary distribution correction estimation, which can handle both discounted and average off-policy evaluation on multiple behavior-agnostic samples.
- 4, TITLE: Principled Weight Initialization for Hypernetworks  
<https://openreview.net/forum?id=H1lma24tPB>  
AUTHORS: Oscar Chang, Lampros Flokas, Hod Lipson  
HIGHLIGHT: The first principled weight initialization method for hypernetworks
- 5, TITLE: On the Convergence of FedAvg on Non-IID Data  
<https://openreview.net/forum?id=HJxNAnVtDS>  
AUTHORS: Xiang Li, Kaixuan Huang, Wenhao Yang, Shusen Wang, Zhihua Zhang  
HIGHLIGHT: The first principled weight initialization method for hypernetworks
- 6, TITLE: Data-dependent Gaussian Prior Objective for Language Generation  
<https://openreview.net/forum?id=S1efxTVYDr>  
AUTHORS: Zuchao Li, Rui Wang, Kehai Chen, Masso Utiyama, Eiichiro Sumita, Zhuosheng Zhang, Hai Zhao  
HIGHLIGHT: We introduce an extra data-dependent Gaussian prior objective to augment the current MLE training, which is designed to capture the prior knowledge in the ground-truth data.
- 7, TITLE: Contrastive Learning of Structured World Models  
<https://openreview.net/forum?id=H1gax6VtDB>  
AUTHORS: Thomas Kipf, Elise van der Pol, Max Welling  
HIGHLIGHT: Contrastively-trained Structured World Models (C-SWMs) learn object-oriented state representations and a relational model of an environment from raw pixel input.
- 8, TITLE: Neural Network Branching for Neural Network Verification  
<https://openreview.net/forum?id=B1evfa4tPB>  
AUTHORS: Jingyue Lu, M. Pawan Kumar  
HIGHLIGHT: We propose a novel learning to branch framework using graph neural networks to improve branch and bound based neural network verification methods.
- 9, TITLE: Why Gradient Clipping Accelerates Training: A Theoretical Justification for Adaptivity  
<https://openreview.net/forum?id=BJgnXpVYwS>  
AUTHORS: Jingzhao Zhang, Tianxing He, Suvrit Sra, Ali Jadbabaie  
HIGHLIGHT: Gradient clipping provably accelerates gradient descent for non-smooth non-convex functions.
- 10, TITLE: Posterior sampling for multi-agent reinforcement learning: solving extensive games with imperfect information  
<https://openreview.net/forum?id=Syg-ET4FPS>  
AUTHORS: Yichi Zhou, Jialian Li, Jun Zhu  
HIGHLIGHT: Gradient clipping provably accelerates gradient descent for non-smooth non-convex functions.
- 11, TITLE: Mogrifier LSTM  
<https://openreview.net/forum?id=SJe5P6EYvS>  
AUTHORS: Gabor Melis, Tom Kocisk, Phil Blunsom  
HIGHLIGHT: An LSTM extension with state-of-the-art language modelling results.

- 12, TITLE: Learning Hierarchical Discrete Linguistic Units from Visually-Grounded Speech  
<https://openreview.net/forum?id=B1eCp4KwH>  
AUTHORS: David Harwath, Wei-Ning Hsu, James Glass  
HIGHLIGHT: Vector quantization layers incorporated into a self-supervised neural model of speech audio learn hierarchical and discrete linguistic units (phone-like, word-like) when trained with a visual-grounding objective.
- 13, TITLE: Mirror-Generative Neural Machine Translation  
<https://openreview.net/forum?id=HkxQRTNYPH>  
AUTHORS: Zaixiang Zheng, Hao Zhou, Shujian Huang, Lei Li, Xin-Yu Dai, Jiajun Chen  
HIGHLIGHT: Vector quantization layers incorporated into a self-supervised neural model of speech audio learn hierarchical and discrete linguistic units (phone-like, word-like) when trained with a visual-grounding objective.
- 14, TITLE: Cyclical Stochastic Gradient MCMC for Bayesian Deep Learning  
<https://openreview.net/forum?id=rkeS1RVtPS>  
AUTHORS: Ruqi Zhang, Chunyuan Li, Jianyi Zhang, Changyou Chen, Andrew Gordon Wilson  
HIGHLIGHT: Vector quantization layers incorporated into a self-supervised neural model of speech audio learn hierarchical and discrete linguistic units (phone-like, word-like) when trained with a visual-grounding objective.
- 15, TITLE: Your classifier is secretly an energy based model and you should treat it like one  
<https://openreview.net/forum?id=Hkxzx0NtDB>  
AUTHORS: Will Grathwohl, Kuan-Chieh Wang, Joern-Henrik Jacobsen, David Duvenaud, Mohammad Norouzi, Kevin Swersky  
HIGHLIGHT: We show that there is a hidden generative model inside of every classifier. We demonstrate how to train this model and show the many benefits of doing so.
- 16, TITLE: Dynamics-Aware Unsupervised Skill Discovery  
<https://openreview.net/forum?id=HJgLR4KvH>  
AUTHORS: Archit Sharma, Shixiang Gu, Sergey Levine, Vikash Kumar, Karol Hausman  
HIGHLIGHT: We propose an unsupervised skill discovery which enables model-based planning for hierarchical reinforcement learning.
- 17, TITLE: Optimal Strategies Against Generative Attacks  
<https://openreview.net/forum?id=BkgzMCVtPB>  
AUTHORS: Roy Mor, Erez Peterfreund, Matan Gavish, Amir Globerson  
HIGHLIGHT: We propose an unsupervised skill discovery which enables model-based planning for hierarchical reinforcement learning.
- 18, TITLE: GraphZoom: A Multi-level Spectral Approach for Accurate and Scalable Graph Embedding  
<https://openreview.net/forum?id=r1lGO0EKDH>  
AUTHORS: Chenhui Deng, Zhiqiang Zhao, Yongyu Wang, Zhiru Zhang, Zhuo Feng  
HIGHLIGHT: A multi-level spectral approach to improving the quality and scalability of unsupervised graph embedding.
- 19, TITLE: Harnessing Structures for Value-Based Planning and Reinforcement Learning  
<https://openreview.net/forum?id=rklHqRVKvH>  
AUTHORS: Yuzhe Yang, Guo Zhang, Zhi Xu, Dina Katabi  
HIGHLIGHT: We propose a generic framework that allows for exploiting the low-rank structure in both planning and deep reinforcement learning.
- 20, TITLE: Comparing Fine-tuning and Rewinding in Neural Network Pruning  
<https://openreview.net/forum?id=S1gSj0NKvB>  
AUTHORS: Alex Renda, Jonathan Frankle, Michael Carbin  
HIGHLIGHT: Instead of fine-tuning after pruning, rewind weights to their values earlier in training and re-train the networks to achieve higher accuracy when pruning neural networks.
- 21, TITLE: Meta-Q-Learning  
<https://openreview.net/forum?id=SJeD3CEFPH>  
AUTHORS: Rasool Fakoor, Pratik Chaudhari, Stefano Soatto, Alexander J. Smola  
HIGHLIGHT: MQL is a simple off-policy meta-RL algorithm that recycles data from the meta-training replay buffer to adapt to new tasks.

- 22, TITLE: Mathematical Reasoning in Latent Space  
<https://openreview.net/forum?id=Ske31kBTPr>  
AUTHORS: Dennis Lee, Christian Szegedy, Markus Rabe, Sarah Loos, Kshitij Bansal  
HIGHLIGHT: Learning to reason about higher order logic formulas in the latent space.
- 23, TITLE: A Theory of Usable Information under Computational Constraints  
<https://openreview.net/forum?id=r1eBeyHFDH>  
AUTHORS: Yilun Xu, Shengjia Zhao, Jiaming Song, Russell Stewart, Stefano Ermon  
HIGHLIGHT: Learning to reason about higher order logic formulas in the latent space.
- 24, TITLE: Geometric Analysis of Nonconvex Optimization Landscapes for Overcomplete Learning  
<https://openreview.net/forum?id=rygixkHKDH>  
AUTHORS: Qing Qu, Yuexiang Zhai, Xiao Li, Yuqian Zhang, Zhihui Zhu  
HIGHLIGHT: Learning to reason about higher order logic formulas in the latent space.
- 25, TITLE: Deep Batch Active Learning by Diverse, Uncertain Gradient Lower Bounds  
<https://openreview.net/forum?id=ryghZJBKPS>  
AUTHORS: Jordan T. Ash, Chicheng Zhang, Akshay Krishnamurthy, John Langford, Alekh Agarwal  
HIGHLIGHT: We introduce a new batch active learning algorithm that's robust to model architecture, batch size, and dataset.
- 26, TITLE: Understanding and Robustifying Differentiable Architecture Search  
<https://openreview.net/forum?id=H1gDNyrKDS>  
AUTHORS: Arber Zela, Thomas Elsken, Tommoy Saikia, Yassine Marrakchi, Thomas Brox, Frank Hutter  
HIGHLIGHT: We study the failure modes of DARTS (Differentiable Architecture Search) by looking at the eigenvalues of the Hessian of validation loss w.r.t. the architecture and propose robustifications based on our analysis.
- 27, TITLE: A Closer Look at Deep Policy Gradients  
<https://openreview.net/forum?id=ryxdEkHtPS>  
AUTHORS: Andrew Ilyas, Logan Engstrom, Shibani Santurkar, Dimitris Tsipras, Firdaus Janoos, Larry Rudolph, Aleksander Madry  
HIGHLIGHT: We study the failure modes of DARTS (Differentiable Architecture Search) by looking at the eigenvalues of the Hessian of validation loss w.r.t. the architecture and propose robustifications based on our analysis.
- 28, TITLE: Implementation Matters in Deep RL: A Case Study on PPO and TRPO  
<https://openreview.net/forum?id=r1etN1rtPB>  
AUTHORS: Logan Engstrom, Andrew Ilyas, Shibani Santurkar, Dimitris Tsipras, Firdaus Janoos, Larry Rudolph, Aleksander Madry  
HIGHLIGHT: We study the failure modes of DARTS (Differentiable Architecture Search) by looking at the eigenvalues of the Hessian of validation loss w.r.t. the architecture and propose robustifications based on our analysis.
- 29, TITLE: Fast Task Inference with Variational Intrinsic Successor Features  
<https://openreview.net/forum?id=BJeAHkrYDS>  
AUTHORS: Steven Hansen, Will Dabney, Andre Barreto, David Warde-Farley, Tom Van de Wiele, Volodymyr Mnih  
HIGHLIGHT: We introduce Variational Intrinsic Successor FeatuRes (VISR), a novel algorithm which learns controllable features that can be leveraged to provide fast task inference through the successor features framework.
- 30, TITLE: Learning to Balance: Bayesian Meta-Learning for Imbalanced and Out-of-distribution Tasks  
<https://openreview.net/forum?id=rkeZIJBYvr>  
AUTHORS: Donghyun Na, Hae Beom Lee, Hayeon Lee, Saehoon Kim, Minseop Park, Eunho Yang, Sung Ju Hwang  
HIGHLIGHT: A novel meta-learning model that adaptively balances the effect of the meta-learning and task-specific learning, and also class-specific learning within each task.
- 31, TITLE: RNA Secondary Structure Prediction By Learning Unrolled Algorithms  
<https://openreview.net/forum?id=S1eALyrYDH>  
AUTHORS: Xinshi Chen, Yu Li, Ramzan Umarov, Xin Gao, Le Song  
HIGHLIGHT: A DL model for RNA secondary structure prediction, which uses an unrolled algorithm in the architecture to enforce constraints.

- 32, TITLE: Watch the Unobserved: A Simple Approach to Parallelizing Monte Carlo Tree Search  
<https://openreview.net/forum?id=BJlQtJlSKDB>  
AUTHORS: Anji Liu, Jianshu Chen, Mingze Yu, Yu Zhai, Xuewen Zhou, Ji Liu  
HIGHLIGHT: We developed an effective parallel UCT algorithm that achieves linear speedup and suffers negligible performance loss.
- 33, TITLE: Target-Embedding Autoencoders for Supervised Representation Learning  
<https://openreview.net/forum?id=BygXfkSYDH>  
AUTHORS: Daniel Jarrett, Mihaela van der Schaar  
HIGHLIGHT: We developed an effective parallel UCT algorithm that achieves linear speedup and suffers negligible performance loss.
- 34, TITLE: Reformer: The Efficient Transformer  
<https://openreview.net/forum?id=rkgNkkHtvB>  
AUTHORS: Nikita Kitaev, Lukasz Kaiser, Anselm Levskaya  
HIGHLIGHT: Efficient Transformer with locality-sensitive hashing and reversible layers
- 35, TITLE: Rotation-invariant clustering of functional cell types in primary visual cortex  
<https://openreview.net/forum?id=rklr9kHFDB>  
AUTHORS: Ivan Ustyuzhaninov, Santiago A. Cadena, Emmanouil Froudarakis, Paul G. Fahey, Edgar Y. Walker, Erick Cobos, Jacob Reimer, Fabian H. Sinz, Andreas S. Tolias, Matthias Bethge, Alexander S. Ecker  
HIGHLIGHT: We classify mouse V1 neurons into putative functional cell types based on their representations in a CNN predicting neural responses
- 36, TITLE: Causal Discovery with Reinforcement Learning  
<https://openreview.net/forum?id=S1g2skStPB>  
AUTHORS: Shengyu Zhu, Ignavier Ng, Zhitang Chen  
HIGHLIGHT: We apply reinforcement learning to score-based causal discovery and achieve promising results on both synthetic and real datasets
- 37, TITLE: Intrinsically Motivated Discovery of Diverse Patterns in Self-Organizing Systems  
<https://openreview.net/forum?id=rkg6sJHYDr>  
AUTHORS: Chris Reinke, Mayalen Etcheverry, Pierre-Yves Oudeyer  
HIGHLIGHT: We study how an unsupervised exploration and feature learning approach addresses efficiently a new problem: automatic discovery of diverse self-organized patterns in high-dim complex systems such as the game of life.
- 38, TITLE: Restricting the Flow: Information Bottlenecks for Attribution  
<https://openreview.net/forum?id=S1xWhlrYwB>  
AUTHORS: Karl Schulz, Leon Sixt, Federico Tombari, Tim Landgraf  
HIGHLIGHT: We apply the informational bottleneck concept to attribution.
- 39, TITLE: Building Deep Equivariant Capsule Networks  
<https://openreview.net/forum?id=BJgNjgSFPS>  
AUTHORS: Sairaam Venkatraman, S. Balasubramanian, R. Raghunatha Sarma  
HIGHLIGHT: A new scalable, group-equivariant model for capsule networks that preserves compositionality under transformations, and is empirically more transformation-robust to older capsule network models.
- 40, TITLE: A Generalized Training Approach for Multiagent Learning  
<https://openreview.net/forum?id=Bkl5kxrKDr>  
AUTHORS: Paul Muller, Shayegan Omidshafiei, Mark Rowland, Karl Tuyls, Julien Perolat, Siqi Liu, Daniel Hennes, Luke Marris, Marc Lanctot, Edward Hughes, Zhe Wang, Guy Lever, Nicolas Heess, Thore Graepel, Remi Munos  
HIGHLIGHT: A new scalable, group-equivariant model for capsule networks that preserves compositionality under transformations, and is empirically more transformation-robust to older capsule network models.
- 41, TITLE: High Fidelity Speech Synthesis with Adversarial Networks  
<https://openreview.net/forum?id=r1gfQgSFDr>  
AUTHORS: Mikolaj Binkowski, Jeff Donahue, Sander Dieleman, Aidan Clark, Erich Elsen, Norman Casagrande, Luis C. Cobo, Karen Simonyan  
HIGHLIGHT: We introduce GAN-TTS, a Generative Adversarial Network for Text-to-Speech, which achieves Mean Opinion Score (MOS) 4.2.

42, TITLE: SEED RL: Scalable and Efficient Deep-RL with Accelerated Central Inference  
<https://openreview.net/forum?id=rkgvXlrKwH>  
AUTHORS: Lasse Espeholt, Raphaël Marinier, Piotr Stanczyk, Ke Wang, Marcin Michalski?  
HIGHLIGHT: SEED RL, a scalable and efficient deep reinforcement learning agent with accelerated central inference. State of the art results, reduces cost and can process millions of frames per second.

43, TITLE: Meta-Learning with Warped Gradient Descent  
<https://openreview.net/forum?id=rkeiQIBFPB>  
AUTHORS: Sebastian Flennerhag, Andrei A. Rusu, Razvan Pascanu, Francesco Visin, Hujun Yin, Raia Hadsell  
HIGHLIGHT: We propose a novel framework for meta-learning a gradient-based update rule that scales to beyond few-shot learning and is applicable to any form of learning, including continual learning.

44, TITLE: Convolutional Conditional Neural Processes  
<https://openreview.net/forum?id=Skey4eBYPS>  
AUTHORS: Jonathan Gordon, Wessel P. Bruinsma, Andrew Y. K. Foong, James Requeima, Yann Dubois, Richard E. Turner  
HIGHLIGHT: We extend deep sets to functional embeddings and Neural Processes to include translation equivariant members

45, TITLE: Gradient Descent Maximizes the Margin of Homogeneous Neural Networks  
<https://openreview.net/forum?id=SJeLIgBKPS>  
AUTHORS: Kaifeng Lyu, Jian Li  
HIGHLIGHT: We study the implicit bias of gradient descent and prove under a minimal set of assumptions that the parameter direction of homogeneous models converges to KKT points of a natural margin maximization problem.

46, TITLE: Adversarial Training and Provable Defenses: Bridging the Gap  
<https://openreview.net/forum?id=SJxSDxrKDr>  
AUTHORS: Mislav Balunovic, Martin Vechev  
HIGHLIGHT: We propose a novel combination of adversarial training and provable defenses which produces a model with state-of-the-art accuracy and certified robustness on CIFAR-10.

47, TITLE: Differentiable Reasoning over a Virtual Knowledge Base  
<https://openreview.net/forum?id=SJxstlHFPH>  
AUTHORS: Bhuvan Dhingra, Manzil Zaheer, Vidhisha Balachandran, Graham Neubig, Ruslan Salakhutdinov, William W. Cohen  
HIGHLIGHT: Differentiable multi-hop access to a textual knowledge base of indexed contextual representations

48, TITLE: Federated Learning with Matched Averaging  
<https://openreview.net/forum?id=BkluqlSFDS>  
AUTHORS: Hongyi Wang, Mikhail Yurochkin, Yuekai Sun, Dimitris Papailiopoulos, Yasaman Khazaeni  
HIGHLIGHT: Communication efficient federated learning with layer-wise matching

49, TITLE: Program Guided Agent  
<https://openreview.net/forum?id=BkxUvnEYDH>  
AUTHORS: Shao-Hua Sun, Te-Lin Wu, Joseph J. Lim  
HIGHLIGHT: We propose a modular framework that can accomplish tasks specified by programs and achieve zero-shot generalization to more complex tasks.

50, TITLE: Sparse Coding with Gated Learned ISTA  
<https://openreview.net/forum?id=BygPO2VKPH>  
AUTHORS: Kailun Wu, Yiwen Guo, Ziang Li, Changshui Zhang  
HIGHLIGHT: We propose gated mechanisms to enhance learned ISTA for sparse coding, with theoretical guarantees on the superiority of the method.

51, TITLE: Graph Neural Networks Exponentially Lose Expressive Power for Node Classification  
<https://openreview.net/forum?id=S1ldO2EFPr>  
AUTHORS: Kenta Oono, Taiji Suzuki  
HIGHLIGHT: We relate the asymptotic behavior of graph neural networks to the graph spectra of underlying graphs and gives principled guidelines for normalizing weights.

- 52, TITLE: Multi-Scale Representation Learning for Spatial Feature Distributions using Grid Cells  
<https://openreview.net/forum?id=rJjdh4KDH>  
AUTHORS: Gengchen Mai, Krzysztof Janowicz, Bo Yan, Rui Zhu, Ling Cai, Ni Lao  
HIGHLIGHT: We propose a representation learning model called Space2vec to encode the absolute positions and spatial relationships of places.
- 53, TITLE: InfoGraph: Unsupervised and Semi-supervised Graph-Level Representation Learning via Mutual Information Maximization  
<https://openreview.net/forum?id=r1Hf2NYvH>  
AUTHORS: Fan-Yun Sun, Jordan Hoffman, Vikas Verma, Jian Tang  
HIGHLIGHT: We propose a representation learning model called Space2vec to encode the absolute positions and spatial relationships of places.
- 54, TITLE: On Robustness of Neural Ordinary Differential Equations  
<https://openreview.net/forum?id=B1e9Y2NYvS>  
AUTHORS: Hanshu YAN, Jiawei DU, Vincent TAN, Jiashi FENG  
HIGHLIGHT: We propose a representation learning model called Space2vec to encode the absolute positions and spatial relationships of places.
- 55, TITLE: Defending Against Physically Realizable Attacks on Image Classification  
<https://openreview.net/forum?id=H1xscnEKDr>  
AUTHORS: Tong Wu, Liang Tong, Yevgeniy Vorobeychik  
HIGHLIGHT: Defending Against Physically Realizable Attacks on Image Classification
- 56, TITLE: Estimating Gradients for Discrete Random Variables by Sampling without Replacement  
<https://openreview.net/forum?id=rkIEj2EFvB>  
AUTHORS: Wouter Kool, Herke van Hoof, Max Welling  
HIGHLIGHT: We derive a low-variance, unbiased gradient estimator for expectations over discrete random variables based on sampling without replacement
- 57, TITLE: Learning to Control PDEs with Differentiable Physics  
<https://openreview.net/forum?id=HyeSin4FPB>  
AUTHORS: Philipp Holl, Nils Thuerey, Vladlen Koltun  
HIGHLIGHT: We train a combination of neural networks to predict optimal trajectories for complex physical systems.
- 58, TITLE: Intensity-Free Learning of Temporal Point Processes  
<https://openreview.net/forum?id=HygOjhEYDH>  
AUTHORS: Oleksandr Shchur, Marin Bilodeau, Stephan G.lemnemmann  
HIGHLIGHT: Learn in temporal point processes by modeling the conditional density, not the conditional intensity.
- 59, TITLE: A Signal Propagation Perspective for Pruning Neural Networks at Initialization  
<https://openreview.net/forum?id=HJeTo2VFwH>  
AUTHORS: Namhoon Lee, Thalaisyasingam Ajanthan, Stephen Gould, Philip H. S. Torr  
HIGHLIGHT: We formally characterize the initialization conditions for effective pruning at initialization and analyze the signal propagation properties of the resulting pruned networks which leads to a method to enhance their trainability and pruning results.
- 60, TITLE: Rethinking the Security of Skip Connections in ResNet-like Neural Networks  
<https://openreview.net/forum?id=BJIRs34Fvr>  
AUTHORS: Dongxian Wu, Yisen Wang, Shu-Tao Xia, James Bailey, Xingjun Ma  
HIGHLIGHT: We identify the security weakness of skip connections in ResNet-like neural networks
- 61, TITLE: WHITE NOISE ANALYSIS OF NEURAL NETWORKS  
<https://openreview.net/forum?id=H1ebhnEYDH>  
AUTHORS: Ali Borji, Sikun Lin  
HIGHLIGHT: We identify the security weakness of skip connections in ResNet-like neural networks
- 62, TITLE: Neural Machine Translation with Universal Visual Representation  
<https://openreview.net/forum?id=Byl8hhNYPS>  
AUTHORS: Zhuosheng Zhang, Kehai Chen, Rui Wang, Masao Utiyama, Eiichiro Sumita, Zuchao Li, Hai Zhao

**HIGHLIGHT:** This work proposed a universal visual representation for neural machine translation (NMT) using retrieved images with similar topics to source sentence, extending image applicability in NMT.

63, **TITLE:** Tranquil Clouds: Neural Networks for Learning Temporally Coherent Features in Point Clouds  
<https://openreview.net/forum?id=BJeKh3VYDH>  
**AUTHORS:** Lukas Prantl, Nuttapong Chentanez, Stefan Jeschke, Nils Thuerey  
**HIGHLIGHT:** We propose a generative neural network approach for temporally coherent point clouds.

64, **TITLE:** PC-DARTS: Partial Channel Connections for Memory-Efficient Architecture Search  
<https://openreview.net/forum?id=BJS634tPr>  
**AUTHORS:** Yuhui Xu, Lingxi Xie, Xiaopeng Zhang, Xin Chen, Guo-Jun Qi, Qi Tian, Hongkai Xiong  
**HIGHLIGHT:** Allowing partial channel connection in super-networks to regularize and accelerate differentiable architecture search

65, **TITLE:** Online and stochastic optimization beyond Lipschitz continuity: A Riemannian approach  
<https://openreview.net/forum?id=rkxZyaNtwB>  
**AUTHORS:** Kimon Antonakopoulos, E. Veronica Belmega, Panayotis Mertikopoulos  
**HIGHLIGHT:** We introduce a novel version of Lipschitz objective continuity that allows stochastic mirror descent methodologies to achieve optimal convergence rates in problems with singularities.

66, **TITLE:** Enhancing Adversarial Defense by k-Winners-Take-All  
<https://openreview.net/forum?id=Skgyv64tvr>  
**AUTHORS:** Chang Xiao, Peilin Zhong, Changxi Zheng  
**HIGHLIGHT:** We propose a simple change to existing neural network structures for better defending against gradient-based adversarial attacks, using the k-winners-take-all activation function.

67, **TITLE:** Encoding word order in complex embeddings  
<https://openreview.net/forum?id=Hke-WTVtwr>  
**AUTHORS:** Benyou Wang, Donghao Zhao, Christina Lioma, Qiuchi Li, Peng Zhang, Jakob Grue Simonsen  
**HIGHLIGHT:** We propose a simple change to existing neural network structures for better defending against gradient-based adversarial attacks, using the k-winners-take-all activation function.

68, **TITLE:** DDSP: Differentiable Digital Signal Processing  
<https://openreview.net/forum?id=B1x1ma4tDr>  
**AUTHORS:** Jesse Engel, Lamtharn (Hanoi) Hantrakul, Chenjie Gu, Adam Roberts  
**HIGHLIGHT:** Better audio synthesis by combining interpretable DSP with end-to-end learning.

69, **TITLE:** Cross-Domain Few-Shot Classification via Learned Feature-Wise Transformation  
<https://openreview.net/forum?id=SJl5Np4tPr>  
**AUTHORS:** Hung-Yu Tseng, Hsin-Ying Lee, Jia-Bin Huang, Ming-Hsuan Yang  
**HIGHLIGHT:** Better audio synthesis by combining interpretable DSP with end-to-end learning.

70, **TITLE:** Ridge Regression: Structure, Cross-Validation, and Sketching  
<https://openreview.net/forum?id=HklRwaEKwB>  
**AUTHORS:** Sifan Liu, Edgar Dobriban  
**HIGHLIGHT:** We study the structure of ridge regression in a high-dimensional asymptotic framework, and get insights about cross-validation and sketching.

71, **TITLE:** Finite Depth and Width Corrections to the Neural Tangent Kernel  
<https://openreview.net/forum?id=SJgndT4KwB>  
**AUTHORS:** Boris Hanin, Mihai Nica  
**HIGHLIGHT:** The neural tangent kernel in a randomly initialized ReLU net is non-trivial fluctuations as long as the depth and width are comparable.

72, **TITLE:** Meta-Learning without Memorization  
<https://openreview.net/forum?id=BklEFpEYwS>  
**AUTHORS:** Mingzhang Yin, George Tucker, Mingyuan Zhou, Sergey Levine, Chelsea Finn  
**HIGHLIGHT:** We identify and formalize the memorization problem in meta-learning and solve this problem with novel meta-regularization method, which greatly expand the domain that meta-learning can be applicable to and effective on.

- 73, TITLE: Influence-Based Multi-Agent Exploration  
<https://openreview.net/forum?id=BJgy96EYvr>  
AUTHORS: Tonghan Wang\*, Jianhao Wang\*, Yi Wu, Chongjie Zhang  
HIGHLIGHT: We identify and formalize the memorization problem in meta-learning and solve this problem with novel meta-regularization method, which greatly expand the domain that meta-learning can be applicable to and effective on.
- 74, TITLE: HOPPITY: LEARNING GRAPH TRANSFORMATIONS TO DETECT AND FIX BUGS IN PROGRAMS  
<https://openreview.net/forum?id=SJeqs6EFvB>  
AUTHORS: Elizabeth Dinella, Hanjun Dai, Ziyang Li, Mayur Naik, Le Song, Ke Wang  
HIGHLIGHT: An learning-based approach for detecting and fixing bugs in Javascript
- 75, TITLE: Sliced Cramer Synaptic Consolidation for Preserving Deeply Learned Representations  
<https://openreview.net/forum?id=BJge3TNKwH>  
AUTHORS: Soheil Kolouri, Nicholas A. Ketz, Andrea Soltoggio, Praveen K. Pilly  
HIGHLIGHT: "A novel framework for overcoming catastrophic forgetting by preserving the distribution of the network's output at an arbitrary layer."
- 76, TITLE: How much Position Information Do Convolutional Neural Networks Encode?  
<https://openreview.net/forum?id=rJeB36NKvB>  
AUTHORS: Md Amirul Islam\*, Sen Jia\*, Neil D. B. Bruce  
HIGHLIGHT: Our work shows positional information has been implicitly encoded in a network. This information is important for detecting position-dependent features, e.g. semantic and saliency.
- 77, TITLE: Hamiltonian Generative Networks  
<https://openreview.net/forum?id=HJenn6VFvB>  
AUTHORS: Aleksandar Botev, Irina Higgins, Andrew Jaegle, Sebastian Racaniere, Danilo J. Rezende, Peter Toth  
HIGHLIGHT: We introduce a class of generative models that reliably learn Hamiltonian dynamics from high-dimensional observations. The learnt Hamiltonian can be applied to sequence modeling or as a normalising flow.
- 78, TITLE: COPHY: Counterfactual Learning of Physical Dynamics  
<https://openreview.net/forum?id=SkeyppEFvS>  
AUTHORS: Fabien Baradel, Natalia Neverova, Julien Mille, Greg Mori, Christian Wolf  
HIGHLIGHT: We introduce a class of generative models that reliably learn Hamiltonian dynamics from high-dimensional observations. The learnt Hamiltonian can be applied to sequence modeling or as a normalising flow.
- 79, TITLE: Estimating counterfactual treatment outcomes over time through adversarially balanced representations  
<https://openreview.net/forum?id=BJg866NFvB>  
AUTHORS: Ioana Bica, Ahmed M Alaa, James Jordon, Mihaela van der Schaar  
HIGHLIGHT: We introduce a class of generative models that reliably learn Hamiltonian dynamics from high-dimensional observations. The learnt Hamiltonian can be applied to sequence modeling or as a normalising flow.
- 80, TITLE: Gradientless Descent: High-Dimensional Zeroth-Order Optimization  
<https://openreview.net/forum?id=Skep6TVYDB>  
AUTHORS: Daniel Golovin, John Karro, Greg Kochanski, Chansoo Lee, Xingyou Song, Qiuyi Zhang  
HIGHLIGHT: Gradientless Descent is a provably efficient gradient-free algorithm that is monotone-invariant and fast for high-dimensional zero-th order optimization.
- 81, TITLE: Conditional Learning of Fair Representations  
<https://openreview.net/forum?id=Hkekl0NFPr>  
AUTHORS: Han Zhao, Amanda Coston, Tameem Adel, Geoffrey J. Gordon  
HIGHLIGHT: We propose a novel algorithm for learning fair representations that can simultaneously mitigate two notions of disparity among different demographic subgroups.
- 82, TITLE: Inductive Matrix Completion Based on Graph Neural Networks  
<https://openreview.net/forum?id=ByxxgCEYDS>  
AUTHORS: Muhan Zhang, Yixin Chen  
HIGHLIGHT: We propose a novel algorithm for learning fair representations that can simultaneously mitigate two notions of disparity among different demographic subgroups.



- 83, TITLE: Duration-of-Stay Storage Assignment under Uncertainty  
<https://openreview.net/forum?id=Hkx7xRVYDr>  
AUTHORS: Michael Lingzhi Li, Elliott Wolf, Daniel Wintz  
HIGHLIGHT: We develop a new storage assignment framework with a novel neural network that enables large efficiency gains in the warehouse.
- 84, TITLE: Emergence of functional and structural properties of the head direction system by optimization of recurrent neural networks  
<https://openreview.net/forum?id=HklSeREtPB>  
AUTHORS: Christopher J. Cueva, Peter Y. Wang, Matthew Chin, Xue-Xin Wei  
HIGHLIGHT: Artificial neural networks trained with gradient descent are capable of recapitulating both realistic neural activity and the anatomical organization of a biological circuit.
- 85, TITLE: Deep neuroethology of a virtual rodent  
<https://openreview.net/forum?id=SyxrxR4KPS>  
AUTHORS: Josh Merel, Diego Aldarondo, Jesse Marshall, Yuval Tassa, Greg Wayne, Bence Olveczky  
HIGHLIGHT: We built a physical simulation of a rodent, trained it to solve a set of tasks, and analyzed the resulting networks.
- 86, TITLE: Doubly Robust Bias Reduction in Infinite Horizon Off-Policy Estimation  
<https://openreview.net/forum?id=S1glGANtDr>  
AUTHORS: Ziyang Tang\*, Yihao Feng\*, Lihong Li, Dengyong Zhou, Qiang Liu  
HIGHLIGHT: We develop a new doubly robust estimator based on the infinite horizon density ratio and off-policy value estimation.
- 87, TITLE: Learning Compositional Koopman Operators for Model-Based Control  
<https://openreview.net/forum?id=H1ldzA4tPr>  
AUTHORS: Yunzhu Li, Hao He, Jiajun Wu, Dina Katabi, Antonio Torralba  
HIGHLIGHT: Learning compositional Koopman operators for efficient system identification and model-based control.
- 88, TITLE: CLEVRER: Collision Events for Video Representation and Reasoning  
<https://openreview.net/forum?id=HkxYzANYDB>  
AUTHORS: Kexin Yi, Chuang Gan, Yunzhu Li, Pushmeet Kohli, Jiajun Wu, Antonio Torralba, Joshua B. Tenenbaum  
HIGHLIGHT: We present a diagnostic dataset for systematic study of temporal and causal reasoning in videos.
- 89, TITLE: The Logical Expressiveness of Graph Neural Networks  
<https://openreview.net/forum?id=r1lZ7AEKvB>  
AUTHORS: Pablo Barceló, Egor V. Kostylev, Mikael Monet, Jorge Pérez, Juan Reutter, Juan Pablo Silva  
HIGHLIGHT: We characterize the expressive power of GNNs in terms of classical logical languages, separating different GNNs and showing connections with standard notions in Knowledge Representation.
- 90, TITLE: The Break-Even Point on the Optimization Trajectories of Deep Neural Networks  
<https://openreview.net/forum?id=r1g87C4KwB>  
AUTHORS: Stanislaw Jastrzebski, Maciej Szymczak, Stanislav Fort, Devansh Arpit, Jacek Tabor, Kyunghyun Cho\*, Krzysztof Geras\*  
HIGHLIGHT: In the early phase of training of deep neural networks there exists a "break-even point" which determines properties of the entire optimization trajectory.
- 91, TITLE: ALBERT: A Lite BERT for Self-supervised Learning of Language Representations  
<https://openreview.net/forum?id=H1eA7AEtvS>  
AUTHORS: Zhenzhong Lan, Mingda Chen, Sebastian Goodman, Kevin Gimpel, Piyush Sharma, Radu Soricut  
HIGHLIGHT: A new pretraining method that establishes new state-of-the-art results on the GLUE, RACE, and SQuAD benchmarks while having fewer parameters compared to BERT-large.
- 92, TITLE: Disentangling neural mechanisms for perceptual grouping  
<https://openreview.net/forum?id=HJxrVA4FDS>  
AUTHORS: Junkyung Kim, Drew Linsley, Kalpit Thakkar, Thomas Serre  
HIGHLIGHT: Horizontal and top-down feedback connections are responsible for complementary perceptual grouping strategies in biological and recurrent vision systems.
- 93, TITLE: Learning to Plan in High Dimensions via Neural Exploration-Exploitation Trees

<https://openreview.net/forum?id=rJgJDAVKvB>  
AUTHORS: Binghong Chen, Bo Dai, Qinjie Lin, Guo Ye, Han Liu, Le Song  
HIGHLIGHT: We propose a meta path planning algorithm which exploits a novel attention-based neural module that can learn generalizable structures from prior experiences to drastically reduce the sample requirement for solving new path planning problems.

94, TITLE: Symplectic Recurrent Neural Networks  
<https://openreview.net/forum?id=BkgYPREtPr>  
AUTHORS: Zhengdao Chen, Jianyu Zhang, Martin Arjovsky, Léon Bottou  
HIGHLIGHT: We propose a meta path planning algorithm which exploits a novel attention-based neural module that can learn generalizable structures from prior experiences to drastically reduce the sample requirement for solving new path planning problems.

95, TITLE: Asymptotics of Wide Networks from Feynman Diagrams  
<https://openreview.net/forum?id=S1gFvANKDS>  
AUTHORS: Ethan Dyer, Guy Gur-Ari  
HIGHLIGHT: A general method for computing the asymptotic behavior of wide networks using Feynman diagrams

96, TITLE: Learning The Difference That Makes A Difference With Counterfactually-Augmented Data  
<https://openreview.net/forum?id=SkIgs0NFvr>  
AUTHORS: Divyansh Kaushik, Eduard Hovy, Zachary Lipton  
HIGHLIGHT: Humans in the loop revise documents to accord with counterfactual labels, resulting resource helps to reduce reliance on spurious associations.  
We will publicly release both datasets.

97, TITLE: Is a Good Representation Sufficient for Sample Efficient Reinforcement Learning?  
<https://openreview.net/forum?id=r1genAVKPB>  
AUTHORS: Simon S. Du, Sham M. Kakade, Ruosong Wang, Lin F. Yang  
HIGHLIGHT: Exponential lower bounds for value-based and policy-based reinforcement learning with function approximation.

98, TITLE: Simplified Action Decoder for Deep Multi-Agent Reinforcement Learning  
<https://openreview.net/forum?id=B1xm3RVtwB>  
AUTHORS: Hengyuan Hu, Jakob N Foerster  
HIGHLIGHT: We develop Simplified Action Decoder, a simple MARL algorithm that beats previous SOTA on Hanabi by a big margin across 2- to 5-player games.

99, TITLE: Network Deconvolution  
<https://openreview.net/forum?id=rkeu30EtvS>  
AUTHORS: Chengxi Ye, Matthew Evanusa, Hua He, Anton Mitrokhin, Thomas Goldstein, James A. Yorke, Cornelia Fermuller, Yiannis Aloimonos  
HIGHLIGHT: We propose a method called network deconvolution that resembles animal vision system to train convolution networks better.

100, TITLE: Neural Symbolic Reader: Scalable Integration of Distributed and Symbolic Representations for Reading Comprehension  
<https://openreview.net/forum?id=ryxjnREFwH>  
AUTHORS: Xinyun Chen, Chen Liang, Adams Wei Yu, Denny Zhou, Dawn Song, Quoc Le  
HIGHLIGHT: We propose a method called network deconvolution that resembles animal vision system to train convolution networks better.

101, TITLE: Real or Not Real, that is the Question  
<https://openreview.net/forum?id=B1lPaCNtPB>  
AUTHORS: Yuanbo Xiangli\*, Yubin Deng\*, Bo Dai\*, Chen Change Loy, Dahua Lin  
HIGHLIGHT: We propose a method called network deconvolution that resembles animal vision system to train convolution networks better.

102, TITLE: Dream to Control: Learning Behaviors by Latent Imagination  
<https://openreview.net/forum?id=S1IOTC4tDS>  
AUTHORS: Danijar Hafner, Timothy Lillicrap, Jimmy Ba, Mohammad Norouzi  
HIGHLIGHT: We present Dreamer, an agent that learns long-horizon behaviors purely by latent imagination using analytic value gradients.

- 103, TITLE: A Probabilistic Formulation of Unsupervised Text Style Transfer  
<https://openreview.net/forum?id=HJIA0C4tPS>  
AUTHORS: Junxian He, Xinyi Wang, Graham Neubig, Taylor Berg-Kirkpatrick  
HIGHLIGHT: We formulate a probabilistic latent sequence model to tackle unsupervised text style transfer, and show its effectiveness across a suite of unsupervised text style transfer tasks.
- 104, TITLE: Emergent Tool Use From Multi-Agent Autocurricula  
<https://openreview.net/forum?id=SkxpxJBKwS>  
AUTHORS: Bowen Baker, Ingmar Kanitscheider, Todor Markov, Yi Wu, Glenn Powell, Bob McGrew, Igor Mordatch  
HIGHLIGHT: We formulate a probabilistic latent sequence model to tackle unsupervised text style transfer, and show its effectiveness across a suite of unsupervised text style transfer tasks.
- 105, TITLE: NAS-Bench-102: Extending the Scope of Reproducible Neural Architecture Search  
<https://openreview.net/forum?id=HJxyZkBKDr>  
AUTHORS: Xuanyi Dong, Yi Yang  
HIGHLIGHT: A NAS benchmark applicable to almost any NAS algorithms.
- 106, TITLE: Strategies for Pre-training Graph Neural Networks  
<https://openreview.net/forum?id=HJIWWJSFDH>  
AUTHORS: Weihua Hu\*, Bowen Liu\*, Joseph Gomes, Marinka Zitnik, Percy Liang, Vijay Pande, Jure Leskovec  
HIGHLIGHT: We develop a strategy for pre-training Graph Neural Networks (GNNs) and systematically study its effectiveness on multiple datasets, GNN architectures, and diverse downstream tasks.
- 107, TITLE: Behaviour Suite for Reinforcement Learning  
<https://openreview.net/forum?id=rygf-kSYwH>  
AUTHORS: Ian Osband, Yotam Doron, Matteo Hessel, John Aslanides, Eren Sezener, Andre Saraiva, Katrina McKinney, Tor Lattimore, Csaba Szepesvari, Satinder Singh, Benjamin Van Roy, Richard Sutton, David Silver, Hado Van Hasselt  
HIGHLIGHT: Bsuite is a collection of carefully-designed experiments that investigate the core capabilities of RL agents.
- 108, TITLE: FreeLB: Enhanced Adversarial Training for Language Understanding  
<https://openreview.net/forum?id=BygzbyHFvB>  
AUTHORS: Chen Zhu, Yu Cheng, Zhe Gan, Siqi Sun, Thomas Goldstein  
HIGHLIGHT: Bsuite is a collection of carefully-designed experiments that investigate the core capabilities of RL agents.
- 109, TITLE: Kernelized Wasserstein Natural Gradient  
<https://openreview.net/forum?id=Hklz71rYvS>  
AUTHORS: M Arbel, A Gretton, W Li, G Montufar  
HIGHLIGHT: Estimator for the Wasserstein natural gradient
- 110, TITLE: And the Bit Goes Down: Revisiting the Quantization of Neural Networks  
<https://openreview.net/forum?id=rJehVyrKwH>  
AUTHORS: Pierre Stock, Armand Joulin, R?mi Gribonval, Benjamin Graham, Herv? J?gou  
HIGHLIGHT: Using a structured quantization technique aiming at better in-domain reconstruction to compress convolutional neural networks
- 111, TITLE: A Latent Morphology Model for Open-Vocabulary Neural Machine Translation  
<https://openreview.net/forum?id=BJxSIISKDH>  
AUTHORS: Duygu Ataman, Wilker Aziz, Alexandra Birch  
HIGHLIGHT: Using a structured quantization technique aiming at better in-domain reconstruction to compress convolutional neural networks
- 112, TITLE: Understanding Why Neural Networks Generalize Well Through GSNR of Parameters  
<https://openreview.net/forum?id=HyevlJStwH>  
AUTHORS: Jinlong Liu, Yunzhi Bai, Guoqing Jiang, Ting Chen, Huayan Wang  
HIGHLIGHT: Using a structured quantization technique aiming at better in-domain reconstruction to compress convolutional neural networks
- 113, TITLE: Model Based Reinforcement Learning for Atari  
<https://openreview.net/forum?id=S1xCPJHtDB>

AUTHORS: Lukasz Kaiser, Mohammad Babaeizadeh, Piotr Milos, Blazej Osinski, Roy H Campbell, Konrad Czechowski, Dumitru Erhan, Chelsea Finn, Piotr Kozakowski, Sergey Levine, Afroz Mohiuddin, Ryan Sepassi, George Tucker, Henryk Michalewski  
HIGHLIGHT: We use video prediction models, a model-based reinforcement learning algorithm and 2h of gameplay per game to train agents for 26 Atari games.

114, TITLE: Disagreement-Regularized Imitation Learning  
<https://openreview.net/forum?id=rkgbYyHtwB>  
AUTHORS: Kianté Brantley, Wen Sun, Mikael Henaff  
HIGHLIGHT: Method for addressing covariate shift in imitation learning using ensemble uncertainty

115, TITLE: Stable Rank Normalization for Improved Generalization in Neural Networks and GANs  
<https://openreview.net/forum?id=H1enKkrFDB>  
AUTHORS: Amartya Sanyal, Philip H. Torr, Puneet K. Dokania  
HIGHLIGHT: We propose Stable Rank Normalisation, a new regulariser based on recent generalization bounds and show how to optimize it with extensive experiments.

116, TITLE: Measuring the Reliability of Reinforcement Learning Algorithms  
<https://openreview.net/forum?id=SJlpYJBKvH>  
AUTHORS: Stephanie C.Y. Chan, Anoop Korattikara, Sam Fishman, John Canny, Sergio Guadarrama  
HIGHLIGHT: A novel set of metrics for measuring reliability of reinforcement learning algorithms (+ accompanying statistical tests)

117, TITLE: Sequential Latent Knowledge Selection for Knowledge-Grounded Dialogue  
<https://openreview.net/forum?id=Hke0K1HKwr>  
AUTHORS: Byeongchang Kim, Jaewoo Ahn, Gunhee Kim  
HIGHLIGHT: Our approach is the first attempt to leverage a sequential latent variable model for knowledge selection in the multi-turn knowledge-grounded dialogue. It achieves the new state-of-the-art performance on Wizard of Wikipedia benchmark.

118, TITLE: Neural Tangents: Fast and Easy Infinite Neural Networks in Python  
<https://openreview.net/forum?id=SkID9yrFPS>  
AUTHORS: Roman Novak, Lechao Xiao, Jiri Hron, Jaehoon Lee, Jascha Sohl-Dickstein, Samuel S. Schoenholz  
HIGHLIGHT: Keras for infinite neural networks.

119, TITLE: Self-labelling via simultaneous clustering and representation learning  
<https://openreview.net/forum?id=Hyx-jyBFPr>  
AUTHORS: Asano YM., Rupprecht C., Vedaldi A.  
HIGHLIGHT: We propose a self-supervised learning formulation that simultaneously learns feature representations and useful dataset labels by optimizing the common cross-entropy loss for features\_and\_labels, while maximizing information.

120, TITLE: The intriguing role of module criticality in the generalization of deep networks  
<https://openreview.net/forum?id=S1e4jkSKvB>  
AUTHORS: Niladri Chatterji, Behnam Neyshabur, Hanie Sedghi  
HIGHLIGHT: We study the phenomenon that some modules of DNNs are more \emph{critical} than others. Our analysis leads us to propose a complexity measure, that is able to explain the superior generalization performance of some architectures over others.

121, TITLE: Harnessing the Power of Infinitely Wide Deep Nets on Small-data Tasks  
<https://openreview.net/forum?id=rkl8sJBYvH>  
AUTHORS: Sanjeev Arora, Simon S. Du, Zhiyuan Li, Ruslan Salakhutdinov, Ruosong Wang, Dingli Yu  
HIGHLIGHT: We verify neural tangent kernel is powerful on small data via experiments on UCI datasets, small CIFAR 10 and low-shot learning on VOC07.

122, TITLE: Differentiation of Blackbox Combinatorial Solvers  
<https://openreview.net/forum?id=BkevoJSYPB>  
AUTHORS: Marin Vlastelica Pogancic, Anselm Paulus, Vit Musil, Georg Martius, Michal Rolinek  
HIGHLIGHT: In this work, we present a method that implements an efficient backward pass through blackbox implementations of combinatorial solvers with linear objective functions.

123, TITLE: Scaling Autoregressive Video Models

- <https://openreview.net/forum?id=rJgsskrFwH>  
AUTHORS: Dirk Weissenborn, Oscar T?ckstr?m, Jakob Uszkoreit  
HIGHLIGHT: We present a novel autoregressive video generation that achieves strong results on popular datasets and produces encouraging continuations of real world videos.
- 124, TITLE: The Ingredients of Real World Robotic Reinforcement Learning  
<https://openreview.net/forum?id=rJe2syrtvS>  
AUTHORS: Henry Zhu, Justin Yu, Abhishek Gupta, Dhruv Shah, Kristian Hartikainen, Avi Singh, Vikash Kumar, Sergey Levine  
HIGHLIGHT: System to learn robotic tasks in the real world with reinforcement learning without instrumentation
- 125, TITLE: Meta-Learning Acquisition Functions for Transfer Learning in Bayesian Optimization  
<https://openreview.net/forum?id=ryeYpJSKwr>  
AUTHORS: Michael Volpp, Lukas Froehlich, Kirsten Fischer, Andreas Doerr, Stefan Falkner, Frank Hutter, Christian Daniel  
HIGHLIGHT: We perform efficient and flexible transfer learning in the framework of Bayesian optimization through meta-learned neural acquisition functions.
- 126, TITLE: Maximum Likelihood Constraint Inference for Inverse Reinforcement Learning  
<https://openreview.net/forum?id=BJliakStvH>  
AUTHORS: Dexter R.R. Scobee, S. Shankar Sastry  
HIGHLIGHT: Our method infers constraints on task execution by leveraging the principle of maximum entropy to quantify how demonstrations differ from expected, un-constrained behavior.
- 127, TITLE: Spectral Embedding of Regularized Block Models  
[https://openreview.net/forum?id=H11\\_0JBYwS](https://openreview.net/forum?id=H11_0JBYwS)  
AUTHORS: Nathan De Lara, Thomas Bonald  
HIGHLIGHT: Graph regularization forces spectral embedding to focus on the largest clusters, making the representation less sensitive to noise.
- 128, TITLE: Towards Hierarchical Importance Attribution: Explaining Compositional Semantics for Neural Sequence Models  
<https://openreview.net/forum?id=BkxRRkSKwr>  
AUTHORS: Xisen Jin, Junyi Du, Zhongyu Wei, Xiangyang Xue, Xiang Ren  
HIGHLIGHT: We propose measurement of phrase importance and algorithms for hierarchical explanation of neural sequence model predictions
- 129, TITLE: word2ket: Space-efficient Word Embeddings inspired by Quantum Entanglement  
<https://openreview.net/forum?id=HkxARkrFwB>  
AUTHORS: Aliakbar Panahi, Seyran Saeedi, Tom Arodz  
HIGHLIGHT: We use ideas from quantum computing to proposed word embeddings that utilize much fewer trainable parameters.
- 130, TITLE: What Can Neural Networks Reason About?  
<https://openreview.net/forum?id=rJxbJeHFPS>  
AUTHORS: Keyulu Xu, Jingling Li, Mozhi Zhang, Simon S. Du, Ken-ichi Kawarabayashi, Stefanie Jegelka  
HIGHLIGHT: We develop a theoretical framework to characterize which reasoning tasks a neural network can learn well.
- 131, TITLE: Training individually fair ML models with sensitive subspace robustness  
<https://openreview.net/forum?id=B1gdkxHFDH>  
AUTHORS: Mikhail Yurochkin, Amanda Bower, Yuekai Sun  
HIGHLIGHT: Algorithm for training individually fair classifier using adversarial robustness
- 132, TITLE: Learning from Rules Generalizing Labeled Exemplars  
<https://openreview.net/forum?id=SkeuexBtDr>  
AUTHORS: Abhijeet Awasthi, Sabyasachi Ghosh, Rasna Goyal, Sunita Sarawagi  
HIGHLIGHT: Coupled rule-exemplar supervision and a implication loss helps to jointly learn to denoise rules and imply labels.
- 133, TITLE: Directional Message Passing for Molecular Graphs

- <https://openreview.net/forum?id=B1eWbxStPH>  
AUTHORS: Johannes Klicpera, Janek Gro?, Stephan G?nnemann  
HIGHLIGHT: Directional message passing incorporates spatial directional information to improve graph neural networks.
- 134, TITLE: Explanation by Progressive Exaggeration  
<https://openreview.net/forum?id=H1xFWgrFPS>  
AUTHORS: Sumedha Singla, Brian Pollack, Junxiang Chen, Kayhan Batmanghelich  
HIGHLIGHT: A method to explain a classifier, by generating visual perturbation of an image by exaggerating or diminishing the semantic features that the classifier associates with a target label.
- 135, TITLE: Compression based bound for non-compressed network: unified generalization error analysis of large compressible deep neural network  
<https://openreview.net/forum?id=ByeGzlrKwH>  
AUTHORS: Taiji Suzuki  
HIGHLIGHT: A method to explain a classifier, by generating visual perturbation of an image by exaggerating or diminishing the semantic features that the classifier associates with a target label.
- 136, TITLE: At Stability's Edge: How to Adjust Hyperparameters to Preserve Minima Selection in Asynchronous Training of Neural Networks?  
<https://openreview.net/forum?id=Bkeb7lHtvH>  
AUTHORS: Niv Giladi, Mor Shpigel Nacson, Elad Hoffer, Daniel Soudry  
HIGHLIGHT: How to prevent stale gradients (in asynchronous SGD) from changing minima stability and degrade steady state generalization?
- 137, TITLE: Disentanglement through Nonlinear ICA with General Incompressible-flow Networks (GIN)  
<https://openreview.net/forum?id=rygeHgSFDH>  
AUTHORS: Peter Sorrenson, Ullrich K?the  
HIGHLIGHT: How to prevent stale gradients (in asynchronous SGD) from changing minima stability and degrade steady state generalization?
- 138, TITLE: Kaleidoscope: An Efficient, Learnable Representation For All Structured Linear Maps  
<https://openreview.net/forum?id=BkgrBgSYDS>  
AUTHORS: Tri Dao, Nimit Sohoni, Albert Gu, Matthew Eichhorn, Amit Blonder, Megan Leszczynski, Atri Rudra  
HIGHLIGHT: We propose a differentiable family of "kaleidoscope matrices," prove that all structured matrices can be represented in this form, and use them to replace hand-crafted linear maps in deep learning models.
- 139, TITLE: Improving Generalization in Meta Reinforcement Learning using Neural Objectives  
<https://openreview.net/forum?id=S1evHerYPr>  
AUTHORS: Louis Kirsch, Sjoerd van Steenkiste, Juergen Schmidhuber  
HIGHLIGHT: We introduce MetaGenRL, a novel meta reinforcement learning algorithm. Unlike prior work, MetaGenRL can generalize to new environments that are entirely different from those used for meta-training.
- 140, TITLE: Drawing Early-Bird Tickets: Toward More Efficient Training of Deep Networks  
<https://openreview.net/forum?id=BJxsrgStvr>  
AUTHORS: Haoran You, Chaojian Li, Pengfei Xu, Yonggan Fu, Yue Wang, Xiaohan Chen, Yingyan Lin, Zhangyang Wang, Richard G. Baraniuk  
HIGHLIGHT: We introduce MetaGenRL, a novel meta reinforcement learning algorithm. Unlike prior work, MetaGenRL can generalize to new environments that are entirely different from those used for meta-training.
- 141, TITLE: Truth or backpropaganda? An empirical investigation of deep learning theory  
<https://openreview.net/forum?id=HyxyIgHFvr>  
AUTHORS: Micah Goldblum, Jonas Geiping, Avi Schwarzschild, Michael Moeller, Tom Goldstein  
HIGHLIGHT: We introduce MetaGenRL, a novel meta reinforcement learning algorithm. Unlike prior work, MetaGenRL can generalize to new environments that are entirely different from those used for meta-training.
- 142, TITLE: Neural Arithmetic Units  
<https://openreview.net/forum?id=H1gNOeHKPS>  
AUTHORS: Andreas Madsen, Alexander Rosenberg Johansen  
HIGHLIGHT: We introduce MetaGenRL, a novel meta reinforcement learning algorithm. Unlike prior work, MetaGenRL can generalize to new environments that are entirely different from those used for meta-training.

143, TITLE: DeepSphere: a graph-based spherical CNN  
<https://openreview.net/forum?id=B1e3OISiPB>  
AUTHORS: Michaël Defferrard, Martino Milani, Frédéric Gusset, Nathanaël Perraudin  
HIGHLIGHT: A graph-based spherical CNN that strikes an interesting balance of trade-offs for a wide variety of applications.

144, TITLE: SUMO: Unbiased Estimation of Log Marginal Probability for Latent Variable Models  
<https://openreview.net/forum?id=SylkYeHtwr>  
AUTHORS: Yucen Luo, Alex Beaton, Mohammad Norouzi, Jun Zhu, David Duvenaud, Ryan P. Adams, Ricky T. Q. Chen  
HIGHLIGHT: We create an unbiased estimator for the log probability of latent variable models, extending such models to a larger scope of applications.

145, TITLE: Deep Learning For Symbolic Mathematics  
<https://openreview.net/forum?id=S1eZYeHFDS>  
AUTHORS: Guillaume Lample, François Charton  
HIGHLIGHT: We train a neural network to compute function integrals, and to solve complex differential equations.

146, TITLE: Making Sense of Reinforcement Learning and Probabilistic Inference  
<https://openreview.net/forum?id=S1xigtHtvS>  
AUTHORS: Brendan O'Donoghue, Ian Osband, Catalin Ionescu  
HIGHLIGHT: Popular algorithms that cast "RL as Inference" ignore the role of uncertainty and exploration. We highlight the importance of these issues and present a coherent framework for RL and inference that handles them gracefully.

147, TITLE: Unbiased Contrastive Divergence Algorithm for Training Energy-Based Latent Variable Models  
<https://openreview.net/forum?id=r1eyceSYPr>  
AUTHORS: Yixuan Qiu, Lingsong Zhang, Xiao Wang  
HIGHLIGHT: We have developed a new training algorithm for energy-based latent variable models that completely removes the bias of contrastive divergence.

148, TITLE: A Mutual Information Maximization Perspective of Language Representation Learning  
<https://openreview.net/forum?id=Syx79eBKwr>  
AUTHORS: Lingpeng Kong, Cyprien de Masson d'Autume, Lei Yu, Wang Ling, Zihang Dai, Dani Yogatama  
HIGHLIGHT: We have developed a new training algorithm for energy-based latent variable models that completely removes the bias of contrastive divergence.

149, TITLE: Energy-based models for atomic-resolution protein conformations  
[https://openreview.net/forum?id=S1e\\_9xrFvS](https://openreview.net/forum?id=S1e_9xrFvS)  
AUTHORS: Yilun Du, Joshua Meier, Jerry Ma, Rob Fergus, Alexander Rives  
HIGHLIGHT: Energy-based models trained on crystallized protein structures predict native side chain configuration and automatically discover molecular energy features.

150, TITLE: Depth-Width Trade-offs for ReLU Networks via Sharkovsky's Theorem  
<https://openreview.net/forum?id=BJe55gBtvH>  
AUTHORS: Vaggos Chatziafratis, Sai Ganesh Nagarajan, Ioannis Panageas, Xiao Wang  
HIGHLIGHT: In this work, we point to a new connection between DNNs expressivity and Sharkovsky's Theorem from dynamical systems, that enables us to characterize the depth-width trade-offs of ReLU networks

151, TITLE: Generalization of Two-layer Neural Networks: An Asymptotic Viewpoint  
<https://openreview.net/forum?id=H1gBsgBYwH>  
AUTHORS: Jimmy Ba, Murat Erdogdu, Taiji Suzuki, Denny Wu, Tianzong Zhang  
HIGHLIGHT: Derived population risk of two-layer neural networks in high dimensions and examined presence / absence of "double descent".

152, TITLE: Reconstructing continuous distributions of 3D protein structure from cryo-EM images  
<https://openreview.net/forum?id=SJxUjIBtwB>  
AUTHORS: Ellen D. Zhong, Tristan Bepler, Joseph H. Davis, Bonnie Berger  
HIGHLIGHT: We propose a deep generative model of volumes for 3D cryo-EM reconstruction from unlabelled 2D images and show that it can learn continuous deformations in protein structure.

153, TITLE: PROGRESSIVE LEARNING AND DISENTANGLEMENT OF HIERARCHICAL REPRESENTATIONS

<https://openreview.net/forum?id=SJxpsxrYPS>

AUTHORS: Zhiyuan Li, Jaideep Vitthal Murkute, Prashna Kumar Gyawali, Linwei Wang

HIGHLIGHT: We proposed a progressive learning method to improve learning and disentangling latent representations at different levels of abstraction.

154, TITLE: AN EXPONENTIAL LEARNING RATE SCHEDULE FOR BATCH NORMALIZED NETWORKS

<https://openreview.net/forum?id=rJg8TeSFDH>

AUTHORS: Zhiyuan Li, Sanjeev Arora

HIGHLIGHT: We propose an exponential learning rate schedule for networks with BatchNorm, which surprisingly performs well in practice and is provably equivalent to popular LR schedules like Step Decay.

155, TITLE: Geom-GCN: Geometric Graph Convolutional Networks

<https://openreview.net/forum?id=S1e2agrFvS>

AUTHORS: Hongbin Pei, Bingzhe Wei, Kevin Chen-Chuan Chang, Yu Lei, Bo Yang

HIGHLIGHT: We propose an exponential learning rate schedule for networks with BatchNorm, which surprisingly performs well in practice and is provably equivalent to popular LR schedules like Step Decay.

156, TITLE: Large Batch Optimization for Deep Learning: Training BERT in 76 minutes

<https://openreview.net/forum?id=Syx4wnEtvH>

AUTHORS: Yang You, Jing Li, Sashank Reddi, Jonathan Hseu, Sanjiv Kumar, Srinadh Bhojanapalli, Xiaodan Song, James Demmel, Kurt Keutzer, Cho-Jui Hsieh

HIGHLIGHT: A fast optimizer for general applications and large-batch training.

157, TITLE: SELF: Learning to Filter Noisy Labels with Self-Ensembling

<https://openreview.net/forum?id=HkgsPhNYPS>

AUTHORS: Duc Tam Nguyen, Chaithanya Kumar Mummadi, Thi Phuong Nhung Ngo, Thi Hoai Phuong Nguyen, Laura Beggel, Thomas Brox

HIGHLIGHT: We propose a self-ensemble framework to train more robust deep learning models under noisy labeled datasets.

158, TITLE: Reinforcement Learning Based Graph-to-Sequence Model for Natural Question Generation

<https://openreview.net/forum?id=HygnDhEtvr>

AUTHORS: Yu Chen, Lingfei Wu, Mohammed J. Zaki

HIGHLIGHT: We propose a self-ensemble framework to train more robust deep learning models under noisy labeled datasets.

159, TITLE: Sharing Knowledge in Multi-Task Deep Reinforcement Learning

<https://openreview.net/forum?id=rkgpv2VFvr>

AUTHORS: Carlo D'Eramo, Davide Tateo, Andrea Bonarini, Marcello Restelli, Jan Peters

HIGHLIGHT: A study on the benefit of sharing representation in Multi-Task Reinforcement Learning.

160, TITLE: On the Weaknesses of Reinforcement Learning for Neural Machine Translation

<https://openreview.net/forum?id=H1eCw3EKvH>

AUTHORS: Leshe Choshen, Lior Fox, Zohar Aizenbud, Omri Abend

HIGHLIGHT: Reinforcement practices for machine translation performance gains might not come from better predictions.

161, TITLE: StructPool: Structured Graph Pooling via Conditional Random Fields

[https://openreview.net/forum?id=BJxg\\_hVtwH](https://openreview.net/forum?id=BJxg_hVtwH)

AUTHORS: Hao Yuan, Shuiwang Ji

HIGHLIGHT: A novel graph pooling method considering relationships between different nodes via conditional random fields.

162, TITLE: Learning deep graph matching with channel-independent embedding and Hungarian attention

<https://openreview.net/forum?id=rJgBd2NYPH>

AUTHORS: Tianshu Yu, Runzhong Wang, Junchi Yan, Baoxin Li

HIGHLIGHT: We proposed a deep graph matching method with novel channel-independent embedding and Hungarian loss, which achieved state-of-the-art performance.

163, TITLE: Graph inference learning for semi-supervised classification

<https://openreview.net/forum?id=r1evOhEKvH>

AUTHORS: Chunyan Xu, Zhen Cui, Xiaobin Hong, Tong Zhang, Jian Yang, Wei Liu

HIGHLIGHT: We propose a novel graph inference learning framework by building structure relations to infer unknown node labels from those labeled nodes in an end-to-end way.



- 164, TITLE: SQL: Imitation Learning via Reinforcement Learning with Sparse Rewards  
<https://openreview.net/forum?id=S1xKd24twB>  
AUTHORS: Siddharth Reddy, Anca D. Dragan, Sergey Levine  
HIGHLIGHT: A simple and effective alternative to adversarial imitation learning: initialize experience replay buffer with demonstrations, set their reward to +1, set reward for all other data to 0, run Q-learning or soft actor-critic to train.
- 165, TITLE: Neural Oblivious Decision Ensembles for Deep Learning on Tabular Data  
<https://openreview.net/forum?id=r1eiu2VtwH>  
AUTHORS: Sergei Popov, Stanislav Morozov, Artem Babenko  
HIGHLIGHT: We propose a new DNN architecture for deep learning on tabular data
- 166, TITLE: Mutual Mean-Teaching: Pseudo Label Refinery for Unsupervised Domain Adaptation on Person Re-identification  
<https://openreview.net/forum?id=rJlnOhVYPS>  
AUTHORS: Yixiao Ge, Dapeng Chen, Hongsheng Li  
HIGHLIGHT: A framework that conducts online refinement of pseudo labels with a novel soft softmax-triplet loss for unsupervised domain adaptation on person re-identification.
- 167, TITLE: Automatically Discovering and Learning New Visual Categories with Ranking Statistics  
[https://openreview.net/forum?id=BJl2\\_nVFPB](https://openreview.net/forum?id=BJl2_nVFPB)  
AUTHORS: Kai Han, Sylvestre-Alvise Rebuffi, Sebastian Ehrhardt, Andrea Vedaldi, Andrew Zisserman  
HIGHLIGHT: A method to automatically discover new categories in unlabelled data, by effectively transferring knowledge from labelled data of other different categories using feature rank statistics.
- 168, TITLE: Maxmin Q-learning: Controlling the Estimation Bias of Q-learning  
<https://openreview.net/forum?id=Bkg0u3Etwr>  
AUTHORS: Qingfeng Lan, Yangchen Pan, Alona Fyshe, Martha White  
HIGHLIGHT: We propose a new variant of Q-learning algorithm called Maxmin Q-learning which provides a parameter-tuning mechanism to flexibly control bias.
- 169, TITLE: Federated Adversarial Domain Adaptation  
<https://openreview.net/forum?id=HJezF3VYPB>  
AUTHORS: Xingchao Peng, Zijun Huang, Yizhe Zhu, Kate Saenko  
HIGHLIGHT: we present a principled approach to the problem of federated domain adaptation, which aims to align the representations learned among the different nodes with the data distribution of the target node.
- 170, TITLE: Depth-Adaptive Transformer  
<https://openreview.net/forum?id=SJg7KhVKPH>  
AUTHORS: Maha Elbayad, Jiatao Gu, Edouard Grave, Michael Auli  
HIGHLIGHT: Sequence model that dynamically adjusts the amount of computation for each input.
- 171, TITLE: DeepHoyer: Learning Sparser Neural Network with Differentiable Scale-Invariant Sparsity Measures  
<https://openreview.net/forum?id=rylBK34FDS>  
AUTHORS: Huanrui Yang, Wei Wen, Hai Li  
HIGHLIGHT: We propose almost everywhere differentiable and scale invariant regularizers for DNN pruning, which can lead to supremum sparsity through standard SGD training.
- 172, TITLE: Evaluating The Search Phase of Neural Architecture Search  
<https://openreview.net/forum?id=H1loF2NFwr>  
AUTHORS: Kaicheng Yu, Christian Sciufo, Martin Jaggi, Claudiu Musat, Mathieu Salzmann  
HIGHLIGHT: We empirically disprove a fundamental hypothesis of the widely-adopted weight sharing strategy in neural architecture search and explain why the state-of-the-arts NAS algorithms performs similarly to random search.
- 173, TITLE: Diverse Trajectory Forecasting with Determinantal Point Processes  
<https://openreview.net/forum?id=ryxnY3NYPS>  
AUTHORS: Ye Yuan, Kris M. Kitani  
HIGHLIGHT: We learn a diversity sampling function with DPPs to obtain a diverse set of samples from a generative model.

- 174, TITLE: Prox-SGD: Training Structured Neural Networks under Regularization and Constraints  
<https://openreview.net/forum?id=HygphEtvR>  
AUTHORS: Yang Yang, Yaxiong Yuan, Avraam Chatzimichailidis, Ruud JG van Sloun, Lei Lei, Symeon Chatzinotas  
HIGHLIGHT: We propose a convergent proximal-type stochastic gradient descent algorithm for constrained nonsmooth nonconvex optimization problems
- 175, TITLE: LAMAL: LAnguage Modeling Is All You Need for Lifelong Language Learning  
<https://openreview.net/forum?id=Skgxcn4YDS>  
AUTHORS: Fan-Keng Sun, Cheng-Hao Ho, Hung-Yi Lee  
HIGHLIGHT: Language modeling is all you need for lifelong language learning.
- 176, TITLE: Learning Expensive Coordination: An Event-Based Deep RL Approach  
<https://openreview.net/forum?id=ryeG924twB>  
AUTHORS: Zhenyu Shi, Runsheng Yu, Xinrun Wang, Rundong Wang, Youzhi Zhang, Hanjiang Lai, Bo An  
HIGHLIGHT: We propose an event-based policy gradient to train the leader and an action abstraction policy gradient to train the followers in leader-follower Markov game.
- 177, TITLE: Curvature Graph Network  
<https://openreview.net/forum?id=BylEqnVFDB>  
AUTHORS: Ze Ye, Kin Sum Liu, Tengfei Ma, Jie Gao, Chao Chen  
HIGHLIGHT: We propose an event-based policy gradient to train the leader and an action abstraction policy gradient to train the followers in leader-follower Markov game.
- 178, TITLE: Distance-Based Learning from Errors for Confidence Calibration  
<https://openreview.net/forum?id=BJeB5hVtvB>  
AUTHORS: Chen Xing, Sercan Arik, Zizhao Zhang, Tomas Pfister  
HIGHLIGHT: We propose an event-based policy gradient to train the leader and an action abstraction policy gradient to train the followers in leader-follower Markov game.
- 179, TITLE: Deep Learning of Determinantal Point Processes via Proper Spectral Sub-gradient  
<https://openreview.net/forum?id=rkelq2VYPr>  
AUTHORS: Tianshu Yu, Yikang Li, Baoxin Li  
HIGHLIGHT: We proposed a specific back-propagation method via proper spectral sub-gradient to integrate determinantal point process to deep learning framework.
- 180, TITLE: N-BEATS: Neural basis expansion analysis for interpretable time series forecasting  
<https://openreview.net/forum?id=r1ecqn4YwB>  
AUTHORS: Boris N. Oreshkin, Dmitri Carпов, Nicolas Chapados, Yoshua Bengio  
HIGHLIGHT: A novel deep interpretable architecture that achieves state of the art on three large scale univariate time series forecasting datasets
- 181, TITLE: Automated Relational Meta-learning  
<https://openreview.net/forum?id=rklp93EtwH>  
AUTHORS: Huaxiu Yao, Xian Wu, Zhiqiang Tao, Yaliang Li, Bolin Ding, Zhenhui Li  
HIGHLIGHT: Addressing task heterogeneity problem in meta-learning by introducing meta-knowledge graph
- 182, TITLE: To Relieve Your Headache of Training an MRF, Take AdvIL  
<https://openreview.net/forum?id=Sylgsn4Fvr>  
AUTHORS: Chongxuan Li, Chao Du, Kun Xu, Max Welling, Jun Zhu, Bo Zhang  
HIGHLIGHT: We propose a black-box algorithm called AdvIL to perform inference and learning on a general Markov random field.
- 183, TITLE: Linear Symmetric Quantization of Neural Networks for Low-precision Integer Hardware  
<https://openreview.net/forum?id=H1lBj2VFPS>  
AUTHORS: Xiandong Zhao, Ying Wang, Xuyi Cai, Cheng Liu, Lei Zhang  
HIGHLIGHT: We introduce an efficient quantization process that allows for performance acceleration on specialized integer-only neural network accelerator.
- 184, TITLE: Weakly Supervised Clustering by Exploiting Unique Class Count  
<https://openreview.net/forum?id=B1xlj3VYvr>

AUTHORS: Mustafa Umit Oner, Hwee Kuan Lee, Wing-Kin Sung  
HIGHLIGHT: A weakly supervised learning based clustering framework performs comparable to that of fully supervised learning models by exploiting unique class count.

185, TITLE: Scalable and Order-robust Continual Learning with Additive Parameter Decomposition  
<https://openreview.net/forum?id=r1gdj2EKPB>  
AUTHORS: Jaehong Yoon, Saehoon Kim, Eunho Yang, Sung Ju Hwang  
HIGHLIGHT: A weakly supervised learning based clustering framework performs comparable to that of fully supervised learning models by exploiting unique class count.

186, TITLE: Continual Learning with Adaptive Weights (CLAW)  
<https://openreview.net/forum?id=Hkls024Kwr>  
AUTHORS: Tameem Adel, Han Zhao, Richard E. Turner  
HIGHLIGHT: A continual learning framework which learns to automatically adapt its architecture based on a proposed variational inference algorithm.

187, TITLE: Transferable Perturbations of Deep Feature Distributions  
<https://openreview.net/forum?id=rJxAo2VYwr>  
AUTHORS: Nathan Inkawhich, Kevin Liang, Lawrence Carin, Yiran Chen  
HIGHLIGHT: We show that perturbations based-on intermediate feature distributions yield more transferable adversarial examples and allow for analysis of the affects of adversarial perturbations on intermediate representations.

188, TITLE: A Learning-based Iterative Method for Solving Vehicle Routing Problems  
<https://openreview.net/forum?id=BJe1334YDH>  
AUTHORS: Hao Lu, Xingwen Zhang, Shuang Yang  
HIGHLIGHT: We show that perturbations based-on intermediate feature distributions yield more transferable adversarial examples and allow for analysis of the affects of adversarial perturbations on intermediate representations.

189, TITLE: Poly-encoders: Architectures and Pre-training Strategies for Fast and Accurate Multi-sentence Scoring  
<https://openreview.net/forum?id=SkxgmnNFvH>  
AUTHORS: Samuel Humeau, Kurt Shuster, Marie-Anne Lachaux, Jason Weston  
HIGHLIGHT: We show that perturbations based-on intermediate feature distributions yield more transferable adversarial examples and allow for analysis of the affects of adversarial perturbations on intermediate representations.

190, TITLE: AutoQ: Automated Kernel-Wise Neural Network Quantization  
<https://openreview.net/forum?id=rygfn4twS>  
AUTHORS: Qian Lou, Feng Guo, Minje Kim, Lantao Liu, Lei Jiang.  
HIGHLIGHT: Accurate, Fast and Automated Kernel-Wise Neural Network Quantization with Mixed Precision using Hierarchical Deep Reinforcement Learning

191, TITLE: Understanding Architectures Learnt by Cell-based Neural Architecture Search  
<https://openreview.net/forum?id=BJxH22EKPS>  
AUTHORS: Yao Shu, Wei Wang, Shaofeng Cai  
HIGHLIGHT: Accurate, Fast and Automated Kernel-Wise Neural Network Quantization with Mixed Precision using Hierarchical Deep Reinforcement Learning

192, TITLE: SVQN: Sequential Variational Soft Q-Learning Networks  
<https://openreview.net/forum?id=r1xPh2VtPB>  
AUTHORS: Shiyu Huang, Hang Su, Jun Zhu, Ting Chen  
HIGHLIGHT: SVQNs formalizes the inference of hidden states and maximum entropy reinforcement learning under a unified graphical model and optimizes the two modules jointly.

193, TITLE: Ranking Policy Gradient  
<https://openreview.net/forum?id=rJld3hEYvS>  
AUTHORS: Kaixiang Lin, Jiayu Zhou  
HIGHLIGHT: We propose ranking policy gradient that learns the optimal rank of actions to maximize return. We propose a general off-policy learning framework with the properties of optimality preserving, variance reduction, and sample-efficiency.

194, TITLE: On Mutual Information Maximization for Representation Learning  
<https://openreview.net/forum?id=rkxoh24FPH>

AUTHORS: Michael Tschannen, Josip Djolonga, Paul K. Rubenstein, Sylvain Gelly, Mario Lucic  
HIGHLIGHT: The success of recent mutual information (MI)-based representation learning approaches strongly depends on the inductive bias in both the choice of network architectures and the parametrization of the employed MI estimators.

195, TITLE: Observational Overfitting in Reinforcement Learning  
<https://openreview.net/forum?id=HJli2hNKDH>  
AUTHORS: Xingyou Song, Yiding Jiang, Yilun Du, Behnam Neyshabur  
HIGHLIGHT: We isolate one factor of RL generalization by analyzing the case when the agent only overfits to the observations. We show that architectural implicit regularizations occur in this regime.

196, TITLE: Enhancing Transformation-Based Defenses Against Adversarial Attacks with a Distribution Classifier  
<https://openreview.net/forum?id=BkgWahEFvr>  
AUTHORS: Connie Kou, Hwee Kuan Lee, Teck Khim Ng, Ee-Chien Chang  
HIGHLIGHT: We enhance existing transformation-based defenses by using a distribution classifier on the distribution of softmax obtained from transformed images.

197, TITLE: Additive Powers-of-Two Quantization: A Non-uniform Discretization for Neural Networks  
<https://openreview.net/forum?id=BkgXT24tDS>  
AUTHORS: Yuhang Li, Xin Dong, Wei Wang  
HIGHLIGHT: We enhance existing transformation-based defenses by using a distribution classifier on the distribution of softmax obtained from transformed images.

198, TITLE: Lazy-CFR: fast and near-optimal regret minimization for extensive games with imperfect information  
<https://openreview.net/forum?id=rJx4p3NYDB>  
AUTHORS: Yichi Zhou, Tongzheng Ren, Jialian Li, Dong Yan, Jun Zhu  
HIGHLIGHT: We enhance existing transformation-based defenses by using a distribution classifier on the distribution of softmax obtained from transformed images.

199, TITLE: Knowledge Consistency between Neural Networks and Beyond  
<https://openreview.net/forum?id=BJeS62Etwh>  
AUTHORS: Ruofan Liang, Tianlin Li, Longfei Li, Quanshi Zhang  
HIGHLIGHT: We enhance existing transformation-based defenses by using a distribution classifier on the distribution of softmax obtained from transformed images.

200, TITLE: Image-guided Neural Object Rendering  
<https://openreview.net/forum?id=Hyg9anEFPS>  
AUTHORS: Justus Thies, Michael Zollhofer, Christian Theobalt, Marc Stamminger, Matthias Nießner  
HIGHLIGHT: We propose a learned image-guided rendering technique that combines the benefits of image-based rendering and GAN-based image synthesis while considering view-dependent effects.

201, TITLE: Implicit Bias of Gradient Descent based Adversarial Training on Separable Data  
<https://openreview.net/forum?id=HkgTTh4FDH>  
AUTHORS: Yan Li, Ethan X. Fang, Huan Xu, Tuo Zhao  
HIGHLIGHT: The solution of gradient descent based adversarial training converges in direction to a robust max margin solution that is adapted to adversary geometry, using L2 perturbation also shows significant speed-up in convergence compared to clean training.

202, TITLE: TabFact: A Large-scale Dataset for Table-based Fact Verification  
<https://openreview.net/forum?id=rkeJRhNYDH>  
AUTHORS: Wenhu Chen, Hongmin Wang, Jianshu Chen, Yunkai Zhang, Hong Wang, Shiyang Li, Xiyu Zhou, William Yang Wang  
HIGHLIGHT: We propose a new dataset to investigate the entailment problem under semi-structured table as premise

203, TITLE: ES-MAML: Simple Hessian-Free Meta Learning  
<https://openreview.net/forum?id=S1exA2NtDB>  
AUTHORS: Xingyou Song, Wenbo Gao, Yuxiang Yang, Krzysztof Choromanski, Aldo Pacchiano, Yunhao Tang  
HIGHLIGHT: We provide a new framework for MAML in the ES/blackbox setting, and show that it allows deterministic and linear policies, better exploration, and non-differentiable adaptation operators.

204, TITLE: Neural Stored-program Memory

- <https://openreview.net/forum?id=rkxxA24FDr>  
AUTHORS: Hung Le, Truyen Tran, Svetha Venkatesh  
HIGHLIGHT: A neural simulation of Universal Turing Machine
- 205, TITLE: Hierarchical Foresight: Self-Supervised Learning of Long-Horizon Tasks via Visual Subgoal Generation  
<https://openreview.net/forum?id=H1gzR2VKDH>  
AUTHORS: Suraj Nair, Chelsea Finn  
HIGHLIGHT: Hierarchical visual foresight learns to generate visual subgoals that break down long-horizon tasks into subtasks, using only self-supervision.
- 206, TITLE: Multi-agent Reinforcement Learning for Networked System Control  
<https://openreview.net/forum?id=Syx7A3NFvH>  
AUTHORS: Tianshu Chu, Sandeep Chinchali, Sachin Katti  
HIGHLIGHT: This paper proposes a new formulation and a new communication protocol for networked multi-agent control problems
- 207, TITLE: FSPool: Learning Set Representations with Featurewise Sort Pooling  
<https://openreview.net/forum?id=HJgBA2VYwH>  
AUTHORS: Yan Zhang, Jonathon Hare, Adam Pr?gel-Bennett  
HIGHLIGHT: Sort in encoder and undo sorting in decoder to avoid responsibility problem in set auto-encoders
- 208, TITLE: Are Pre-trained Language Models Aware of Phrases? Simple but Strong Baselines for Grammar Induction  
<https://openreview.net/forum?id=H1xPR3NtPB>  
AUTHORS: Taek Kim, Jihun Choi, Daniel Edmiston, Sang-goo Lee  
HIGHLIGHT: Sort in encoder and undo sorting in decoder to avoid responsibility problem in set auto-encoders
- 209, TITLE: Dynamically Pruned Message Passing Networks for Large-scale Knowledge Graph Reasoning  
<https://openreview.net/forum?id=rkeuAhVKvB>  
AUTHORS: Xiaoran Xu, Wei Feng, Yunsheng Jiang, Xiaohui Xie, Zhiqing Sun, Zhi-Hong Deng  
HIGHLIGHT: We propose to learn an input-dependent subgraph, dynamically and selectively expanded, to explicitly model a sequential reasoning process.
- 210, TITLE: Mixup Inference: Better Exploiting Mixup to Defend Adversarial Attacks  
<https://openreview.net/forum?id=ByxtC2VtPB>  
AUTHORS: Tianyu Pang\*, Kun Xu\*, Jun Zhu  
HIGHLIGHT: We exploit the global linearity of the mixup-trained models in inference to break the locality of the adversarial perturbations.
- 211, TITLE: Theory and Evaluation Metrics for Learning Disentangled Representations  
<https://openreview.net/forum?id=HJgK0h4Ywr>  
AUTHORS: Kien Do, Truyen Tran  
HIGHLIGHT: We exploit the global linearity of the mixup-trained models in inference to break the locality of the adversarial perturbations.
- 212, TITLE: Measuring Compositional Generalization: A Comprehensive Method on Realistic Data  
<https://openreview.net/forum?id=SygcCnNKwr>  
AUTHORS: Daniel Keysers, Nathanael Sch?rli, Nathan Scales, Hylke Buisman, Daniel Furrer, Sergii Kashubin, Nikola Momchev, Danila Sinopalnikov, Lukasz Stafiniak, Tibor Tihon, Dmitry Tsarkov, Xiao Wang, Marc van Zee, Olivier Bousquet  
HIGHLIGHT: Benchmark and method to measure compositional generalization by maximizing divergence of compound frequency at small divergence of atom frequency.
- 213, TITLE: Rethinking Softmax Cross-Entropy Loss for Adversarial Robustness  
<https://openreview.net/forum?id=Byg9A24tvB>  
AUTHORS: Tianyu Pang, Kun Xu, Yinpeng Dong, Chao Du, Ning Chen, Jun Zhu  
HIGHLIGHT: Applying the softmax function in training leads to indirect and unexpected supervision on features. We propose a new training objective to explicitly induce dense feature regions for locally sufficient samples to benefit adversarial robustness.
- 214, TITLE: The Implicit Bias of Depth: How Incremental Learning Drives Generalization  
<https://openreview.net/forum?id=H1lj0nNFwB>  
AUTHORS: Daniel Gissin, Shai Shalev-Shwartz, Amit Daniely

- HIGHLIGHT: We study the sparsity-inducing bias of deep models, caused by their learning dynamics.
- 215, TITLE: The Variational Bandwidth Bottleneck: Stochastic Evaluation on an Information Budget  
<https://openreview.net/forum?id=Hye1kTVFDS>  
AUTHORS: Anirudh Goyal, Yoshua Bengio, Matthew Botvinick, Sergey Levine  
HIGHLIGHT: Training agents with adaptive computation based on information bottleneck can promote generalization.
- 216, TITLE: Learning the Arrow of Time for Problems in Reinforcement Learning  
<https://openreview.net/forum?id=rylJkpEtwS>  
AUTHORS: Nasim Rahaman, Steffen Wolf, Anirudh Goyal, Roman Remme, Yoshua Bengio  
HIGHLIGHT: We learn the arrow of time for MDPs and use it to measure reachability, detect side-effects and obtain a curiosity reward signal.
- 217, TITLE: Reinforcement Learning with Competitive Ensembles of Information-Constrained Primitives  
<https://openreview.net/forum?id=ryxgJTEYDr>  
AUTHORS: Anirudh Goyal, Shagun Sodhani, Jonathan Binas, Xue Bin Peng, Sergey Levine, Yoshua Bengio  
HIGHLIGHT: Learning an implicit master policy, as a master policy in HRL can fail to generalize.
- 218, TITLE: Robust Local Features for Improving the Generalization of Adversarial Training  
<https://openreview.net/forum?id=H1lZJpVFvr>  
AUTHORS: Chuanbiao Song, Kun He, Jiadong Lin, Liwei Wang, John E. Hopcroft  
HIGHLIGHT: We propose a new stream of adversarial training approach called Robust Local Features for Adversarial Training (RLFAT) that significantly improves both the adversarially robust generalization and the standard generalization.
- 219, TITLE: Analysis of Video Feature Learning in Two-Stream CNNs on the Example of Zebrafish Swim Bout Classification  
<https://openreview.net/forum?id=rJgQkT4twH>  
AUTHORS: Bennet Breier, Arno Onken  
HIGHLIGHT: We demonstrate the utility of a recent AI explainability technique by visualizing the learned features of a CNN trained on binary classification of zebrafish movements.
- 220, TITLE: Learning Disentangled Representations for Counterfactual Regression  
<https://openreview.net/forum?id=HkxBJT4YvB>  
AUTHORS: Negar Hassanpour, Russell Greiner  
HIGHLIGHT: We demonstrate the utility of a recent AI explainability technique by visualizing the learned features of a CNN trained on binary classification of zebrafish movements.
- 221, TITLE: Exploration in Reinforcement Learning with Deep Covering Options  
<https://openreview.net/forum?id=SkeIyaVtwB>  
AUTHORS: Yuu Jinnai, Jee Won Park, Marlos C. Machado, George Konidaris  
HIGHLIGHT: We introduce a method to automatically discover task-agnostic options that encourage exploration for reinforcement learning.
- 222, TITLE: AE-OT: A NEW GENERATIVE MODEL BASED ON EXTENDED SEMI-DISCRETE OPTIMAL TRANSPORT  
<https://openreview.net/forum?id=HklDyTNYwH>  
AUTHORS: Dongsheng An, Yang Guo, Na Lei, Zhongxuan Luo, Shing-Tung Yau, Xianfeng Gu  
HIGHLIGHT: We introduce a method to automatically discover task-agnostic options that encourage exploration for reinforcement learning.
- 223, TITLE: Logic and the 2-Simplicial Transformer  
<https://openreview.net/forum?id=rkecJ6VFvr>  
AUTHORS: James Clift, Dmitry Doryn, Daniel Murfet, James Wallbridge  
HIGHLIGHT: We introduce the 2-simplicial Transformer and show that this architecture is a useful inductive bias for logical reasoning in the context of deep reinforcement learning.
- 224, TITLE: Watch, Try, Learn: Meta-Learning from Demonstrations and Rewards  
<https://openreview.net/forum?id=SJg5J6NtDr>  
AUTHORS: Allan Zhou, Eric Jang, Daniel Kappler, Alex Herzog, Mohi Khansari, Paul Wohlhart, Yunfei Bai, Mrinal Kalakrishnan, Sergey Levine, Chelsea Finn

**HIGHLIGHT:** We introduce the 2-simplicial Transformer and show that this architecture is a useful inductive bias for logical reasoning in the context of deep reinforcement learning.

225, **TITLE:** Fooling Detection Alone is Not Enough: Adversarial Attack against Multiple Object Tracking  
<https://openreview.net/forum?id=rJl31TNYPr>  
**AUTHORS:** Yunhan Jia, Yantao Lu, Junjie Shen, Qi Alfred Chen, Hao Chen, Zhenyu Zhong, Tao Wei  
**HIGHLIGHT:** We study the adversarial machine learning attacks against the Multiple Object Tracking mechanisms for the first time.

226, **TITLE:** DivideMix: Learning with Noisy Labels as Semi-supervised Learning  
<https://openreview.net/forum?id=HJgExaVtwr>  
**AUTHORS:** Junnan Li, Steven C.H. Hoi, Richard Socher  
**HIGHLIGHT:** We propose a novel framework for learning with noisy labels by leveraging semi-supervised learning.

227, **TITLE:** Improving Adversarial Robustness Requires Revisiting Misclassified Examples  
<https://openreview.net/forum?id=rklOg6EFwS>  
**AUTHORS:** Yisen Wang, Difan Zou, Jinfeng Yi, James Bailey, Xingjun Ma, Quanquan Gu  
**HIGHLIGHT:** By differentiating misclassified and correctly classified data, we propose a new misclassification aware defense that improves the state-of-the-art adversarial robustness.

228, **TITLE:** V-MPO: On-Policy Maximum a Posteriori Policy Optimization for Discrete and Continuous Control  
<https://openreview.net/forum?id=SylOlP4FvH>  
**AUTHORS:** H. Francis Song, Abbas Abdolmaleki, Jost Tobias Springenberg, Aidan Clark, Hubert Soyer, Jack W. Rae, Seb Noury, Arun Ahuja, Siqi Liu, Dhruva Tirumala, Nicolas Heess, Dan Belov, Martin Riedmiller, Matthew M. Botvinick  
**HIGHLIGHT:** A state-value function-based version of MPO that achieves good results in a wide range of tasks in discrete and continuous control.

229, **TITLE:** Attributes Obfuscation with Complex-Valued Features  
<https://openreview.net/forum?id=S1xF164tDr>  
**AUTHORS:** Liyao Xiang, Hao Zhang, Haotian Ma, Yifan Zhang, Jie Ren, Quanshi Zhang  
**HIGHLIGHT:** A state-value function-based version of MPO that achieves good results in a wide range of tasks in discrete and continuous control.

230, **TITLE:** Accelerating SGD with momentum for over-parameterized learning  
<https://openreview.net/forum?id=r1gixp4FPH>  
**AUTHORS:** Chaoyue Liu, Mikhail Belkin  
**HIGHLIGHT:** This work proves the non-acceleration of Nesterov SGD with any hyper-parameters, and proposes new algorithm which provably accelerates SGD in the over-parameterized setting.

231, **TITLE:** A critical analysis of self-supervision, or what we can learn from a single image  
<https://openreview.net/forum?id=B1esx6EYvr>  
**AUTHORS:** Asano YM., Rupprecht C., Vedaldi A.  
**HIGHLIGHT:** We evaluate self-supervised feature learning methods and find that with sufficient data augmentation early layers can be learned using just one image. This is informative about self-supervision and the role of augmentations.

232, **TITLE:** Disentangling Factors of Variations Using Few Labels  
<https://openreview.net/forum?id=SygagpEKwB>  
**AUTHORS:** Francesco Locatello, Michael Tschannen, Stefan Bauer, Gunnar R?tsch, Bernhard Sch?lkopf, Olivier Bachem  
**HIGHLIGHT:** We evaluate self-supervised feature learning methods and find that with sufficient data augmentation early layers can be learned using just one image. This is informative about self-supervision and the role of augmentations.

233, **TITLE:** Functional vs. parametric equivalence of ReLU networks  
<https://openreview.net/forum?id=Bylx-TNKvH>  
**AUTHORS:** Mary Phuong, Christoph H. Lampert  
**HIGHLIGHT:** We prove that there exist ReLU networks whose parameters are almost uniquely determined by the function they implement.

234, **TITLE:** Input Complexity and Out-of-distribution Detection with Likelihood-based Generative Models  
<https://openreview.net/forum?id=SyxIWpVYvr>  
**AUTHORS:** Joan Serr?, David ?lvarez, Vicen? G?mez, Olga Slizovskaia, Jos? F. N??ez, Jordi Luque

**HIGHLIGHT:** We pose that generative models' likelihoods are excessively influenced by the input's complexity, and propose a way to compensate it when detecting out-of-distribution inputs

235, **TITLE:** RTFM: Generalising to New Environment Dynamics via Reading

<https://openreview.net/forum?id=SJgob6NKvH>

**AUTHORS:** Victor Zhong, Tim Rockt?schel, Edward Grefenstette

**HIGHLIGHT:** We show language understanding via reading is promising way to learn policies that generalise to new environments.

236, **TITLE:** What graph neural networks cannot learn: depth vs width

<https://openreview.net/forum?id=B1l2bp4YwS>

**AUTHORS:** Andreas Loukas

**HIGHLIGHT:** Several graph problems are impossible unless the product of a graph neural network's depth and width exceeds (a function of) the graph size.

237, **TITLE:** Progressive Memory Banks for Incremental Domain Adaptation

<https://openreview.net/forum?id=BkepbnNFwr>

**AUTHORS:** Nabihha Asghar, Lili Mou, Kira A. Selby, Kevin D. Pantasdo, Pascal Poupart, Xin Jiang

**HIGHLIGHT:** We present a neural memory-based architecture for incremental domain adaptation, and provide theoretical and empirical results.

238, **TITLE:** Automated curriculum generation through setter-solver interactions

<https://openreview.net/forum?id=H1e0Wp4KvH>

**AUTHORS:** Andrew Lampinen, Sebastien Racaniere, Adam Santoro, David Reichert, Vlad Firoiu, Timothy Lillicrap

**HIGHLIGHT:** We investigate automatic curriculum generation and identify a number of losses useful to learn to generate a curriculum of tasks.

239, **TITLE:** On Identifiability in Transformers

<https://openreview.net/forum?id=BJglf6EFDB>

**AUTHORS:** Gino Brunner, Yang Liu, Damian Pascual, Oliver Richter, Massimiliano Ciaramita, Roger Wattenhofer

**HIGHLIGHT:** We investigate the identifiability and interpretability of attention distributions and tokens within contextual embeddings in the self-attention based BERT model.

240, **TITLE:** Exploring Model-based Planning with Policy Networks

<https://openreview.net/forum?id=H1exf64KwH>

**AUTHORS:** Tingwu Wang, Jimmy Ba

**HIGHLIGHT:** how to achieve state-of-the-art performance by combining policy network in model-based planning

241, **TITLE:** Learning Self-Correctable Policies and Value Functions from Demonstrations with Negative Sampling

<https://openreview.net/forum?id=rke-f6NKvS>

**AUTHORS:** Yuping Luo, Huazhe Xu, Tengyu Ma

**HIGHLIGHT:** We introduce a notion of conservatively-extrapolated value functions, which provably lead to policies that can self-correct to stay close to the demonstration states, and learn them with a novel negative sampling technique.

242, **TITLE:** Geometric Insights into the Convergence of Nonlinear TD Learning

<https://openreview.net/forum?id=SJezGp4YPr>

**AUTHORS:** David Brandfonbrener, Joan Bruna

**HIGHLIGHT:** We introduce a notion of conservatively-extrapolated value functions, which provably lead to policies that can self-correct to stay close to the demonstration states, and learn them with a novel negative sampling technique.

243, **TITLE:** Few-shot Text Classification with Distributional Signatures

<https://openreview.net/forum?id=H1emfT4twB>

**AUTHORS:** Yujia Bao, Menghua Wu, Shiyu Chang, Regina Barzilay

**HIGHLIGHT:** Meta-learning methods used for vision, directly applied to NLP, perform worse than nearest neighbors on new classes; we can do better with distributional signatures.

244, **TITLE:** Escaping Saddle Points Faster with Stochastic Momentum

<https://openreview.net/forum?id=rkeNfp4tPr>

**AUTHORS:** Jun-Kun Wang, Chi-Heng Lin, Jacob Abernethy

**HIGHLIGHT:** Higher momentum parameter  $\beta$  helps for escaping saddle points faster



- 245, TITLE: Adversarial Policies: Attacking Deep Reinforcement Learning  
<https://openreview.net/forum?id=HJgEMpVFwB>  
AUTHORS: Adam Gleave, Michael Dennis, Cody Wild, Neel Kant, Sergey Levine, Stuart Russell  
HIGHLIGHT: Deep RL policies can be attacked by other agents taking actions so as to create natural observations that are adversarial.
- 246, TITLE: VideoFlow: A Conditional Flow-Based Model for Stochastic Video Generation  
<https://openreview.net/forum?id=rJgUfTEYvH>  
AUTHORS: Manoj Kumar, Mohammad Babaeizadeh, Dumitru Erhan, Chelsea Finn, Sergey Levine, Laurent Dinh, Durk Kingma  
HIGHLIGHT: We demonstrate that flow-based generative models offer a viable and competitive approach to generative modeling of video.
- 247, TITLE: GLAD: Learning Sparse Graph Recovery  
<https://openreview.net/forum?id=BkxpMTEtPB>  
AUTHORS: Harsh Shrivastava, Xinshi Chen, Binghong Chen, Guanghui Lan, Srinivas Aluru, Han Liu, Le Song  
HIGHLIGHT: A data-driven learning algorithm based on unrolling the Alternating Minimization optimization for sparse graph recovery.
- 248, TITLE: Pruned Graph Scattering Transforms  
<https://openreview.net/forum?id=rJeg7TEYwB>  
AUTHORS: Vassilis N. Ioannidis, Siheng Chen, Georgios B. Giannakis  
HIGHLIGHT: A data-driven learning algorithm based on unrolling the Alternating Minimization optimization for sparse graph recovery.
- 249, TITLE: Pretrained Encyclopedia: Weakly Supervised Knowledge-Pretrained Language Model  
<https://openreview.net/forum?id=BJlzm64tDH>  
AUTHORS: Wenhan Xiong, Jingfei Du, William Yang Wang, Veselin Stoyanov  
HIGHLIGHT: A data-driven learning algorithm based on unrolling the Alternating Minimization optimization for sparse graph recovery.
- 250, TITLE: Can gradient clipping mitigate label noise?  
<https://openreview.net/forum?id=rklB76EKPr>  
AUTHORS: Aditya Krishna Menon, Ankit Singh Rawat, Sashank J. Reddi, Sanjiv Kumar  
HIGHLIGHT: Gradient clipping doesn't endow robustness to label noise, but a simple loss-based variant does.
- 251, TITLE: Editable Neural Networks  
<https://openreview.net/forum?id=HJedXaEtvS>  
AUTHORS: Anton Sinitsin, Vsevolod Plokhotnyuk, Dmitry Pyrkun, Sergei Popov, Artem Babenko  
HIGHLIGHT: Training neural networks so you can efficiently patch them later.
- 252, TITLE: LEARNING EXECUTION THROUGH NEURAL CODE FUSION  
<https://openreview.net/forum?id=SJetQpEYvB>  
AUTHORS: Zhan Shi, Kevin Swersky, Daniel Tarlow, Parthasarathy Ranganathan, Milad Hashemi  
HIGHLIGHT: Training neural networks so you can efficiently patch them later.
- 253, TITLE: FasterSeg: Searching for Faster Real-time Semantic Segmentation  
<https://openreview.net/forum?id=BJgqQ6NYvB>  
AUTHORS: Wuyang Chen, Xinyu Gong, Xianming Liu, Qian Zhang, Yuan Li, Zhangyang Wang  
HIGHLIGHT: We present a real-time segmentation model automatically discovered by a multi-scale NAS framework, achieving 30% faster than state-of-the-art models.
- 254, TITLE: Difference-Seeking Generative Adversarial Network--Unseen Sample Generation  
<https://openreview.net/forum?id=rygimpVFvB>  
AUTHORS: Yi Lin Sung, Sung-Hsien Hsieh, Soo-Chang Pei, Chun-Shien Lu  
HIGHLIGHT: We proposed a novel GAN framework to generate unseen data.
- 255, TITLE: Stochastic AUC Maximization with Deep Neural Networks

- <https://openreview.net/forum?id=HJepXaVYDr>  
AUTHORS: Mingrui Liu, Zhuoning Yuan, Yiming Ying, Tianbao Yang  
HIGHLIGHT: The paper designs two algorithms for the stochastic AUC maximization problem with state-of-the-art complexities when using deep neural network as predictive model, which are also verified by empirical studies.
- 256, TITLE: Semantically-Guided Representation Learning for Self-Supervised Monocular Depth  
<https://openreview.net/forum?id=ByxT7TNFvH>  
AUTHORS: Vitor Guizilini, Rui Hou, Jie Li, Rares Ambrus, Adrien Gaidon  
HIGHLIGHT: We propose a novel semantically-guided architecture for self-supervised monocular depth estimation
- 257, TITLE: MACER: Attack-free and Scalable Robust Training via Maximizing Certified Radius  
<https://openreview.net/forum?id=rJx1Na4Fwr>  
AUTHORS: Runtian Zhai, Chen Dan, Di He, Huan Zhang, Boqing Gong, Pradeep Ravikumar, Cho-Jui Hsieh, Liwei Wang  
HIGHLIGHT: We propose MACER: a provable defense algorithm that trains robust models by maximizing the certified radius. It does not use adversarial training but performs better than all existing provable l2-defenses.
- 258, TITLE: Detecting and Diagnosing Adversarial Images with Class-Conditional Capsule Reconstructions  
<https://openreview.net/forum?id=Skgy464Kvr>  
AUTHORS: Yao Qin, Nicholas Frosst, Sara Sabour, Colin Raffel, Garrison Cottrell, Geoffrey Hinton  
HIGHLIGHT: We propose MACER: a provable defense algorithm that trains robust models by maximizing the certified radius. It does not use adversarial training but performs better than all existing provable l2-defenses.
- 259, TITLE: Adversarial Example Detection and Classification with Asymmetrical Adversarial Training  
<https://openreview.net/forum?id=SJeQEp4YDH>  
AUTHORS: Xuwang Yin, Soheil Kolouri, Gustavo K Rohde  
HIGHLIGHT: A new generative modeling technique based on asymmetrical adversarial training, and its applications to adversarial example detection and robust classification
- 260, TITLE: Variational Recurrent Models for Solving Partially Observable Control Tasks  
<https://openreview.net/forum?id=r1lL4a4tDB>  
AUTHORS: Dongqi Han, Kenji Doya, Jun Tani  
HIGHLIGHT: A deep RL algorithm for solving POMDPs by auto-encoding the underlying states using a variational recurrent model
- 261, TITLE: Population-Guided Parallel Policy Search for Reinforcement Learning  
<https://openreview.net/forum?id=rJeINp4KwH>  
AUTHORS: Whiyoung Jung, Giseung Park, Youngchul Sung  
HIGHLIGHT: A deep RL algorithm for solving POMDPs by auto-encoding the underlying states using a variational recurrent model
- 262, TITLE: Compositional languages emerge in a neural iterated learning model  
<https://openreview.net/forum?id=HkePNpVKPB>  
AUTHORS: Yi Ren, Shangmin Guo, Matthieu Labeau, Shay B. Cohen, Simon Kirby  
HIGHLIGHT: Use iterated learning framework to facilitate the dominance of high compositional language in multi-agent games.
- 263, TITLE: Black-Box Adversarial Attack with Transferable Model-based Embedding  
<https://openreview.net/forum?id=SJxhNTNYwB>  
AUTHORS: Zhichao Huang, Tong Zhang  
HIGHLIGHT: We present a new method that combines transfer-based and scored black-box adversarial attack, improving the success rate and query efficiency of black-box adversarial attack across different network architectures.
- 264, TITLE: I Am Going MAD: Maximum Discrepancy Competition for Comparing Classifiers Adaptively  
<https://openreview.net/forum?id=rJehNT4YPr>  
AUTHORS: Haotao Wang, Tianlong Chen, Zhangyang Wang, Kede Ma  
HIGHLIGHT: We present an efficient and adaptive framework for comparing image classifiers to maximize the discrepancies between the classifiers, in place of comparing on fixed test sets.
- 265, TITLE: Mixout: Effective Regularization to Finetune Large-scale Pretrained Language Models  
<https://openreview.net/forum?id=HkgaETNtDB>

AUTHORS: Cheolhyoung Lee, Kyunghyun Cho, Wanmo Kang  
HIGHLIGHT: We present an efficient and adaptive framework for comparing image classifiers to maximize the discrepancies between the classifiers, in place of comparing on fixed test sets.

266, TITLE: Q-learning with UCB Exploration is Sample Efficient for Infinite-Horizon MDP  
<https://openreview.net/forum?id=BkgISTNFDB>  
AUTHORS: Yuanhao Wang, Kefan Dong, Xiaoyu Chen, Liwei Wang  
HIGHLIGHT: We adapt Q-learning with UCB-exploration bonus to infinite-horizon MDP with discounted rewards without accessing a generative model, and improves the previously best known result.

267, TITLE: Deep Network classification by Scattering and Homotopy dictionary learning  
<https://openreview.net/forum?id=SJxWS64FwH>  
AUTHORS: John Zarka, Louis Thiry, Tomas Angles, Stephane Mallat  
HIGHLIGHT: A scattering transform followed by supervised dictionary learning reaches a higher accuracy than AlexNet on ImageNet.

268, TITLE: Data-Independent Neural Pruning via Coresets  
<https://openreview.net/forum?id=H1gmHaEKwB>  
AUTHORS: Ben Mussay, Margarita Osadchy, Vladimir Braverman, Samson Zhou, Dan Feldman  
HIGHLIGHT: We propose an efficient, provable and data independent method for network compression via neural pruning using coresets of neurons -- a novel construction proposed in this paper.

269, TITLE: Bounds on Over-Parameterization for Guaranteed Existence of Descent Paths in Shallow ReLU Networks  
<https://openreview.net/forum?id=BkgXHTNtvS>  
AUTHORS: Arsalan Sharifnassab, Saber Salehkaleybar, S. Jamaloddin Golestani  
HIGHLIGHT: We propose an efficient, provable and data independent method for network compression via neural pruning using coresets of neurons -- a novel construction proposed in this paper.

270, TITLE: Novelty Detection Via Blurring  
<https://openreview.net/forum?id=ByeNra4FDB>  
AUTHORS: Sungik Choi, Sae-Young Chung  
HIGHLIGHT: We propose a novel OOD detector that employ blurred images as adversarial examples . Our model achieve significant OOD detection performance in various domains.

271, TITLE: Nonlinearities in activations substantially shape the loss surfaces of neural networks  
<https://openreview.net/forum?id=B1x6BTEKwr>  
AUTHORS: Fengxiang He, Bohan Wang, Dacheng Tao  
HIGHLIGHT: This paper presents how the loss surfaces of nonlinear neural networks are substantially shaped by the nonlinearities in activations.

272, TITLE: Relational State-Space Model for Stochastic Multi-Object Systems  
<https://openreview.net/forum?id=B1IGU64tDr>  
AUTHORS: Fan Yang, Ling Chen, Fan Zhou, Yusong Gao, Wei Cao  
HIGHLIGHT: A deep hierarchical state-space model in which the state transitions of correlated objects are coordinated by graph neural networks.

273, TITLE: Learning Efficient Parameter Server Synchronization Policies for Distributed SGD  
<https://openreview.net/forum?id=rJxX8T4Kvr>  
AUTHORS: Rong Zhu, Sheng Yang, Andreas Pfadler, Zhengping Qian, Jingren Zhou  
HIGHLIGHT: We apply a reinforcement learning based approach to learning optimal synchronization policies used for Parameter Server-based distributed training of SGD.

274, TITLE: Action Semantics Network: Considering the Effects of Actions in Multiagent Systems  
<https://openreview.net/forum?id=ryg48p4tPH>  
AUTHORS: Weixun Wang, Tianpei Yang, Yong Liu, Jianye Hao, Xiaotian Hao, Yujing Hu, Yingfeng Chen, Changjie Fan, Yang Gao  
HIGHLIGHT: Our proposed ASN characterizes different actions' influence on other agents using neural networks based on the action semantics between them.

275, TITLE: Vid2Game: Controllable Characters Extracted from Real-World Videos

<https://openreview.net/forum?id=SkxBUpEKwH>  
AUTHORS: Oran Gafni, Lior Wolf, Yaniv Taigman  
HIGHLIGHT: Our proposed ASN characterizes different actions' influence on other agents using neural networks based on the action semantics between them.

276, TITLE: Self-Adversarial Learning with Comparative Discrimination for Text Generation  
<https://openreview.net/forum?id=B1I8L6EtDS>  
AUTHORS: Wangchunshu Zhou, Tao Ge, Ke Xu, Furu Wei, Ming Zhou  
HIGHLIGHT: We propose a self-adversarial learning (SAL) paradigm which improves the generator in a self-play fashion for improving GANs' performance in text generation.

277, TITLE: Robust training with ensemble consensus  
<https://openreview.net/forum?id=ryxOUTVYDH>  
AUTHORS: Jisoo Lee, Sae-Young Chung  
HIGHLIGHT: This work presents a method of generating and using ensembles effectively to identify noisy examples in the presence of annotation noise.

278, TITLE: Identifying through Flows for Recovering Latent Representations  
<https://openreview.net/forum?id=SkIOUpEYvB>  
AUTHORS: Shen Li, Bryan Hooi, Gim Hee Lee  
HIGHLIGHT: This work presents a method of generating and using ensembles effectively to identify noisy examples in the presence of annotation noise.

279, TITLE: Certified Robustness for Top-k Predictions against Adversarial Perturbations via Randomized Smoothing  
<https://openreview.net/forum?id=BkeWw6VFwr>  
AUTHORS: Jinyuan Jia, Xiaoyu Cao, Binghui Wang, Neil Zhenqiang Gong  
HIGHLIGHT: We study the certified robustness for top-k predictions via randomized smoothing under Gaussian noise and derive a tight robustness bound in  $L_2$  norm.

280, TITLE: Optimistic Exploration even with a Pessimistic Initialisation  
<https://openreview.net/forum?id=r1xGP6VYwH>  
AUTHORS: Tabish Rashid, Bei Peng, Wendelin Boehmer, Shimon Whiteson  
HIGHLIGHT: We augment the Q-value estimates with a count-based bonus that ensures optimism during action selection and bootstrapping, even if the Q-value estimates are pessimistic.

281, TITLE: VL-BERT: Pre-training of Generic Visual-Linguistic Representations  
<https://openreview.net/forum?id=SygXPaEYvH>  
AUTHORS: Weijie Su, Xizhou Zhu, Yue Cao, Bin Li, Lewei Lu, Furu Wei, Jifeng Dai  
HIGHLIGHT: VL-BERT is a simple yet powerful pre-trainable generic representation for visual-linguistic tasks. It is pre-trained on the massive-scale caption dataset and text-only corpus, and can be finetuned for various down-stream visual-linguistic tasks.

282, TITLE: Deformable Kernels: Adapting Effective Receptive Fields for Object Deformation  
<https://openreview.net/forum?id=SkxSv6VFvS>  
AUTHORS: Hang Gao, Xizhou Zhu, Stephen Lin, Jifeng Dai  
HIGHLIGHT: Don't deform your convolutions -- deform your kernels.

283, TITLE: Ensemble Distribution Distillation  
<https://openreview.net/forum?id=BygSP6Vtvr>  
AUTHORS: Andrey Malinin, Bruno Mlodozienec, Mark Gales  
HIGHLIGHT: We distill an ensemble of models into a single model, capturing both the improved classification performance and information about the diversity of the ensemble, which is useful for uncertainty estimation.

284, TITLE: Gap-Aware Mitigation of Gradient Staleness  
<https://openreview.net/forum?id=B1ILw6EYwB>  
AUTHORS: Saar Barkai, Ido Hakimi, Assaf Schuster  
HIGHLIGHT: A new distributed, asynchronous, SGD-based algorithm, which achieves state-of-the-art accuracy on existing architectures using staleness penalization without having to re-tune the hyperparameters.

285, TITLE: Counterfactuals uncover the modular structure of deep generative models  
<https://openreview.net/forum?id=SJxDDpEKvH>

AUTHORS: Michel Besserve, Arash Mehrjou, Remy Sun, Bernhard Schoelkopf  
HIGHLIGHT: We develop a framework to find modular internal representations in generative models and manipulate them to generate counterfactual examples.

286, TITLE: Physics-as-Inverse-Graphics: Unsupervised Physical Parameter Estimation from Video  
<https://openreview.net/forum?id=BJeKwTNFvB>  
AUTHORS: Miguel Jaques, Michael Burke, Timothy Hospedales  
HIGHLIGHT: We propose a model that is able to perform physical parameter estimation of systems from video, where the differential equations governing the scene dynamics are known, but labeled states or objects are not available.

287, TITLE: An Inductive Bias for Distances: Neural Nets that Respect the Triangle Inequality  
<https://openreview.net/forum?id=HJeiDpVFPr>  
AUTHORS: Silviu Pitis, Harris Chan, Kiarash Jamali, Jimmy Ba  
HIGHLIGHT: We propose novel neural network architectures, guaranteed to satisfy the triangle inequality, for purposes of (asymmetric) metric learning and modeling graph distances.

288, TITLE: A Constructive Prediction of the Generalization Error Across Scales  
<https://openreview.net/forum?id=ryenvpEKDr>  
AUTHORS: Jonathan S. Rosenfeld, Amir Rosenfeld, Yonatan Belinkov, Nir Shavit  
HIGHLIGHT: We predict the generalization error and specify the model which attains it across model/data scales.

289, TITLE: Scalable Neural Methods for Reasoning With a Symbolic Knowledge Base  
<https://openreview.net/forum?id=BJlguT4YPr>  
AUTHORS: William W. Cohen, Haitian Sun, R. Alex Hofer, Matthew Siegler  
HIGHLIGHT: A scalable differentiable neural module that implements reasoning on symbolic KBs.

290, TITLE: CLN2INV: Learning Loop Invariants with Continuous Logic Networks  
<https://openreview.net/forum?id=HJlfuTEtvB>  
AUTHORS: Gabriel Ryan, Justin Wong, Jianan Yao, Ronghui Gu, Suman Jana  
HIGHLIGHT: We introduce the Continuous Logic Network (CLN), a novel neural architecture for automatically learning loop invariants and general SMT formulas.

291, TITLE: NAS evaluation is frustratingly hard  
<https://openreview.net/forum?id=HygrdpVKvr>  
AUTHORS: Antoine Yang, Pedro M. Esperan?a, Fabio M. Carlucci  
HIGHLIGHT: A study of how different components in the NAS pipeline contribute to the final accuracy. Also, a benchmark of 8 methods on 5 datasets.

292, TITLE: Efficient and Information-Preserving Future Frame Prediction and Beyond  
[https://openreview.net/forum?id=B1eY\\_pVYvB](https://openreview.net/forum?id=B1eY_pVYvB)  
AUTHORS: Wei Yu, Yichao Lu, Steve Easterbrook, Sanja Fidler  
HIGHLIGHT: A study of how different components in the NAS pipeline contribute to the final accuracy. Also, a benchmark of 8 methods on 5 datasets.

293, TITLE: Order Learning and Its Application to Age Estimation  
<https://openreview.net/forum?id=HygsuaNFwr>  
AUTHORS: Kyungsun Lim, Nyeong-Ho Shin, Young-Yoon Lee, Chang-Su Kim  
HIGHLIGHT: The notion of order learning is proposed and it is applied to regression problems in computer vision

294, TITLE: ReClor: A Reading Comprehension Dataset Requiring Logical Reasoning  
<https://openreview.net/forum?id=HJgJtT4tvB>  
AUTHORS: Weihao Yu, Zihang Jiang, Yanfei Dong, Jiashi Feng  
HIGHLIGHT: We introduce ReClor, a reading comprehension dataset requiring logical reasoning, and find that current state-of-the-art models struggle with real logical reasoning with poor performance near that of random guess.

295, TITLE: AssembleNet: Searching for Multi-Stream Neural Connectivity in Video Architectures  
<https://openreview.net/forum?id=SJgMK64Ywr>  
AUTHORS: Michael S. Ryoo, AJ Piergiovanni, Mingxing Tan, Anelia Angelova  
HIGHLIGHT: We search for multi-stream neural architectures with better connectivity and spatio-temporal interactions for video understanding.

- 296, TITLE: Adversarially Robust Representations with Smooth Encoders  
<https://openreview.net/forum?id=H1gfFaEYDS>  
AUTHORS: Taylan Cemgil, Sumedh Ghaisas, Krishnamurthy (Dj) Dvijotham, Pushmeet Kohli  
HIGHLIGHT: We propose a method for computing adversarially robust representations in an entirely unsupervised way.
- 297, TITLE: From Variational to Deterministic Autoencoders  
<https://openreview.net/forum?id=S1g7tpEYDS>  
AUTHORS: Partha Ghosh, Mehdi S. M. Sajjadi, Antonio Vergari, Michael Black, Bernhard Scholkopf  
HIGHLIGHT: Deterministic regularized autoencoders can learn a smooth, meaningful latent space as VAEs without having to force some arbitrarily chosen prior (i.e., Gaussian).
- 298, TITLE: Computation Reallocation for Object Detection  
<https://openreview.net/forum?id=SkxLFaNKwB>  
AUTHORS: Feng Liang, Ronghao Guo, Chen Lin, Ming Sun, Wei Wu, Junjie Yan, Wanli Ouyang  
HIGHLIGHT: We propose CR-NAS to reallocate engaged computation resources in different resolution and spatial position.
- 299, TITLE: Finding and Visualizing Weaknesses of Deep Reinforcement Learning Agents  
<https://openreview.net/forum?id=rylvYaNYDH>  
AUTHORS: Christian Rupprecht, Cyril Ibrahim, Christopher J. Pal  
HIGHLIGHT: We generate critical states of a trained RL algorithms to visualize potential weaknesses.
- 300, TITLE: A Fair Comparison of Graph Neural Networks for Graph Classification  
<https://openreview.net/forum?id=HygDF6NFPB>  
AUTHORS: Federico Errica, Marco Podda, Davide Bacciu, Alessio Micheli  
HIGHLIGHT: We provide a rigorous comparison of different Graph Neural Networks for graph classification.
- 301, TITLE: Size-free generalization bounds for convolutional neural networks  
[https://openreview.net/forum?id=r1e\\_FpNFDr](https://openreview.net/forum?id=r1e_FpNFDr)  
AUTHORS: Philip M. Long, Hanie Sedghi  
HIGHLIGHT: We prove generalization bounds for convolutional neural networks that take account of weight-tying
- 302, TITLE: SAdam: A Variant of Adam for Strongly Convex Functions  
<https://openreview.net/forum?id=rye5YaEtPr>  
AUTHORS: Guanghui Wang, Shiyin Lu, Quan Cheng, Weiwei Tu, Lijun Zhang  
HIGHLIGHT: A variant of Adam for strongly convex functions
- 303, TITLE: Continual Learning with Bayesian Neural Networks for Non-Stationary Data  
<https://openreview.net/forum?id=SJIsFpVtDB>  
AUTHORS: Richard Kurle, Botond Cseke, Alexej Klushyn, Patrick van der Smagt, Stephan G?nnemann  
HIGHLIGHT: This work addresses continual learning for non-stationary data, using Bayesian neural networks and memory-based online variational Bayes.
- 304, TITLE: Multiplicative Interactions and Where to Find Them  
<https://openreview.net/forum?id=rylnK6VtDH>  
AUTHORS: Siddhant M. Jayakumar, Jacob Menick, Wojciech M. Czarnecki, Jonathan Schwarz, Jack Rae, Simon Osindero, Yee Whye Teh, Tim Harley, Razvan Pascanu  
HIGHLIGHT: We explore the role of multiplicative interaction as a unifying framework to describe a range of classical and modern neural network architectural motifs, such as gating, attention layers, hypernetworks, and dynamic convolutions amongst others.
- 305, TITLE: FEW-SHOT LEARNING ON GRAPHS VIA SUPER-CLASSES BASED ON GRAPH SPECTRAL MEASURES  
<https://openreview.net/forum?id=Bkeeca4Kvr>  
AUTHORS: Jatin Chauhan, Deepak Nathani, Manohar Kaul  
HIGHLIGHT: We explore the role of multiplicative interaction as a unifying framework to describe a range of classical and modern neural network architectural motifs, such as gating, attention layers, hypernetworks, and dynamic convolutions amongst others.

- 306, TITLE: ON COMPUTATION AND GENERALIZATION OF GENERATIVE ADVERSARIAL IMITATION LEARNING  
https://openreview.net/forum?id=BJl-5pNKDB  
AUTHORS: Minshuo Chen, Yizhou Wang, Tianyi Liu, Zhuoran Yang, Xingguo Li, Zhaoran Wang, Tuo Zhao  
HIGHLIGHT: We explore the role of multiplicative interaction as a unifying framework to describe a range of classical and modern neural network architectural motifs, such as gating, attention layers, hypernetworks, and dynamic convolutions amongst others.
- 307, TITLE: A TARGET-AGNOSTIC ATTACK ON DEEP MODELS: EXPLOITING SECURITY VULNERABILITIES OF TRANSFER LEARNING  
https://openreview.net/forum?id=BylVcTNtDS  
AUTHORS: Shahbaz Rezaei, Xin Liu  
HIGHLIGHT: We explore the role of multiplicative interaction as a unifying framework to describe a range of classical and modern neural network architectural motifs, such as gating, attention layers, hypernetworks, and dynamic convolutions amongst others.
- 308, TITLE: Low-Resource Knowledge-Grounded Dialogue Generation  
https://openreview.net/forum?id=rJelcTNtvS  
AUTHORS: Xueliang Zhao, Wei Wu, Chongyang Tao, Can Xu, Dongyan Zhao, Rui Yan  
HIGHLIGHT: We explore the role of multiplicative interaction as a unifying framework to describe a range of classical and modern neural network architectural motifs, such as gating, attention layers, hypernetworks, and dynamic convolutions amongst others.
- 309, TITLE: Deep 3D Pan via Local adaptive "t-shaped" convolutions with global and local adaptive dilations  
https://openreview.net/forum?id=BlgF56VYPH  
AUTHORS: Juan Luis Gonzalez Bello, Munchurl Kim  
HIGHLIGHT: Novel architecture for stereoscopic view synthesis at arbitrary camera shifts utilizing adaptive t-shaped kernels with adaptive dilations.
- 310, TITLE: Tree-Structured Attention with Hierarchical Accumulation  
https://openreview.net/forum?id=HJxK5pEYvr  
AUTHORS: Xuan-Phi Nguyen, Shafiq Joty  
HIGHLIGHT: Novel architecture for stereoscopic view synthesis at arbitrary camera shifts utilizing adaptive t-shaped kernels with adaptive dilations.
- 311, TITLE: The asymptotic spectrum of the Hessian of DNN throughout training  
https://openreview.net/forum?id=SkgscaNYPS  
AUTHORS: Arthur Jacot, Franck Gabriel, Clement Hongler  
HIGHLIGHT: Description of the limiting spectrum of the Hessian of the loss surface of DNNs in the infinite-width limit.
- 312, TITLE: Actor-Critic Provably Finds Nash Equilibria of Linear-Quadratic Mean-Field Games  
https://openreview.net/forum?id=H1lhqpeYPr  
AUTHORS: Zuyue Fu, Zhuoran Yang, Yongxin Chen, Zhaoran Wang  
HIGHLIGHT: Actor-Critic method with function approximation finds the Nash equilibrium pairs in mean-field games with theoretical guarantee.
- 313, TITLE: In Search for a SAT-friendly Binarized Neural Network Architecture  
https://openreview.net/forum?id=SJx-j64FDr  
AUTHORS: Nina Narodytska, Hongce Zhang, Aarti Gupta, Toby Walsh  
HIGHLIGHT: Formal analysis of Binarized Neural Networks
- 314, TITLE: Generative Ratio Matching Networks  
https://openreview.net/forum?id=SJg7spEYDS  
AUTHORS: Akash Srivastava, Kai Xu, Michael U. Gutmann, Charles Sutton  
HIGHLIGHT: Formal analysis of Binarized Neural Networks
- 315, TITLE: Learning to Represent Programs with Property Signatures  
https://openreview.net/forum?id=rylHspEKPr  
AUTHORS: Augustus Odena, Charles Sutton  
HIGHLIGHT: We represent a computer program using a set of simpler programs and use this representation to improve program synthesis techniques.

- 316, TITLE: V4D: 4D Covolutional Neural Networks for Video-level Representations Learning  
<https://openreview.net/forum?id=SJeLopEYDH>  
AUTHORS: Shiwen Zhang, Sheng Guo, Weilin Huang, Matthew R. Scott, Limin Wang  
HIGHLIGHT: A novel 4D CNN structure for video-level representation learning, surpassing recent 3D CNNs.
- 317, TITLE: Option Discovery using Deep Skill Chaining  
<https://openreview.net/forum?id=B1gqipNYwH>  
AUTHORS: Akhil Bagaria, George Konidakis  
HIGHLIGHT: We present a new hierarchical reinforcement learning algorithm which can solve high-dimensional goal-oriented tasks far more reliably than non-hierarchical agents and other state-of-the-art skill discovery techniques.
- 318, TITLE: Quantifying the Cost of Reliable Photo Authentication via High-Performance Learned Lossy Representations  
<https://openreview.net/forum?id=HyxG3p4twS>  
AUTHORS: Pawel Korus, Nasir Memon  
HIGHLIGHT: We learn an efficient lossy image compression codec which can be optimized to facilitate reliable photo manipulation detection at fractional cost in payload/quality and even at low bitrates.
- 319, TITLE: On the Variance of the Adaptive Learning Rate and Beyond  
<https://openreview.net/forum?id=rkgz2aEKDr>  
AUTHORS: Liyuan Liu, Haoming Jiang, Pengcheng He, Weizhu Chen, Xiaodong Liu, Jianfeng Gao, Jiawei Han  
HIGHLIGHT: If warmup is the answer, what is the question?
- 320, TITLE: Dynamical Distance Learning for Semi-Supervised and Unsupervised Skill Discovery  
<https://openreview.net/forum?id=H1lmhaVtvr>  
AUTHORS: Kristian Hartikainen, Xinyang Geng, Tuomas Haamoja, Sergey Levine  
HIGHLIGHT: We show how to automatically learn dynamical distances in reinforcement learning setting and use them to provide well-shaped reward functions for reaching new goals.
- 321, TITLE: A Theoretical Analysis of the Number of Shots in Few-Shot Learning  
<https://openreview.net/forum?id=HkgB2TNYPs>  
AUTHORS: Tianshi Cao, Marc T Law, Sanja Fidler  
HIGHLIGHT: The paper analyzes the effect of shot number on prototypical networks and proposes a robust method when the shot number differs from meta-training to meta-testing time.
- 322, TITLE: Unsupervised Model Selection for Variational Disentangled Representation Learning  
<https://openreview.net/forum?id=SyxL2Tntvr>  
AUTHORS: Sunny Duan, Loic Matthey, Andre Saraiva, Nick Watters, Chris Burgess, Alexander Lerchner, Irina Higgins  
HIGHLIGHT: We introduce a method for unsupervised disentangled model selection for VAE-based disentangled representation learning approaches.
- 323, TITLE: Extracting and Leveraging Feature Interaction Interpretations  
<https://openreview.net/forum?id=BkgnhTetDS>  
AUTHORS: Michael Tsang, Dehua Cheng, Hanpeng Liu, Xue Feng, Eric Zhou, Yan Liu  
HIGHLIGHT: Proposed a method to extract and leverage interpretations of feature interactions
- 324, TITLE: Understanding the Limitations of Variational Mutual Information Estimators  
<https://openreview.net/forum?id=B1x62TntDS>  
AUTHORS: Jiaming Song, Stefano Ermon  
HIGHLIGHT: Proposed a method to extract and leverage interpretations of feature interactions
- 325, TITLE: GENESIS: Generative Scene Inference and Sampling with Object-Centric Latent Representations  
<https://openreview.net/forum?id=BkxfaTVFwH>  
AUTHORS: Martin Engelcke, Adam R. Kosiorek, Oiwi Parker Jones, Ingmar Posner  
HIGHLIGHT: We present the first object-centric generative model of 3D visual scenes capable of both decomposing and generating scenes.
- 326, TITLE: Language GANs Falling Short  
<https://openreview.net/forum?id=BJgza6VtPB>



AUTHORS: Massimo Caccia, Lucas Caccia, William Fedus, Hugo Larochelle, Joelle Pineau, Laurent Charlin  
HIGHLIGHT: GANs have been applied to text generation and are believed SOTA. However, we propose a new evaluation protocol demonstrating that maximum-likelihood trained models are still better.

327, TITLE: Stochastic Conditional Generative Networks with Basis Decomposition  
<https://openreview.net/forum?id=S11SapVtwS>  
AUTHORS: Ze Wang, Xiuyuan Cheng, Guillermo Sapiro, Qiang Qiu  
HIGHLIGHT: GANs have been applied to text generation and are believed SOTA. However, we propose a new evaluation protocol demonstrating that maximum-likelihood trained models are still better.

328, TITLE: LEARNED STEP SIZE QUANTIZATION  
<https://openreview.net/forum?id=rkgO66VKDS>  
AUTHORS: Steven K. Esser, Jeffrey L. McKinstry, Deepika Bablani, Rathinakumar Appuswamy, Dharmendra S. Modha  
HIGHLIGHT: A method for learning quantization configuration for low precision networks that achieves state of the art performance for quantized networks.

329, TITLE: On the "steerability" of generative adversarial networks  
<https://openreview.net/forum?id=HylsTT4FvB>  
AUTHORS: Ali Jahanian\*, Lucy Chai\*, Phillip Isola  
HIGHLIGHT: A method for learning quantization configuration for low precision networks that achieves state of the art performance for quantized networks.

330, TITLE: Reinforced active learning for image segmentation  
<https://openreview.net/forum?id=SkG6TNFvr>  
AUTHORS: Arantxa Casanova, Pedro O. Pinheiro, Negar Rostamzadeh, Christopher J. Pal  
HIGHLIGHT: Learning a labeling policy with reinforcement learning to reduce labeling effort for the task of semantic segmentation

331, TITLE: Sign Bits Are All You Need for Black-Box Attacks  
<https://openreview.net/forum?id=SygW0TEFwH>  
AUTHORS: Abdullah Al-Dujaili, Una-May O'Reilly  
HIGHLIGHT: We present a sign-based, rather than magnitude-based, gradient estimation approach that shifts gradient estimation from continuous to binary black-box optimization.

332, TITLE: Deep Semi-Supervised Anomaly Detection  
<https://openreview.net/forum?id=HkgH0TEYwH>  
AUTHORS: Lukas Ruff, Robert A. Vandermeulen, Nico G?rnitz, Alexander Binder, Emmanuel M?ller, Klaus-Robert M?ller, Marius Kloft  
HIGHLIGHT: We introduce Deep SAD, a deep method for general semi-supervised anomaly detection that especially takes advantage of labeled anomalies.

333, TITLE: Budgeted Training: Rethinking Deep Neural Network Training Under Resource Constraints  
<https://openreview.net/forum?id=HyxLRTVKPH>  
AUTHORS: Mengtian Li, Ersin Yumer, Deva Ramanan  
HIGHLIGHT: Introduce a formal setting for budgeted training and propose a budget-aware linear learning rate schedule

334, TITLE: Minimizing FLOPs to Learn Efficient Sparse Representations  
<https://openreview.net/forum?id=SygpC6Ntvr>  
AUTHORS: Biswajit Paria, Chih-Kuan Yeh, Ning Xu, Barnabas Poczos, Pradeep Ravikumar, Ian E.H. Yen  
HIGHLIGHT: We propose an approach to learn sparse high dimensional representations that are fast to search, by incorporating a surrogate of the number of operations directly into the loss function.

335, TITLE: Reanalysis of Variance Reduced Temporal Difference Learning  
<https://openreview.net/forum?id=S1ly10EKDS>  
AUTHORS: Tengyu Xu, Zhe Wang, Yi Zhou, Yingbin Liang  
HIGHLIGHT: This paper provides a rigorous study of the variance reduced TD learning and characterizes its advantage over vanilla TD learning

336, TITLE: Imitation Learning via Off-Policy Distribution Matching  
<https://openreview.net/forum?id=Hyg-JC4FDr>

AUTHORS: Ilya Kostrikov, Ofir Nachum, Jonathan Tompson  
HIGHLIGHT: This paper provides a rigorous study of the variance reduced TD learning and characterizes its advantage over vanilla TD learning

337, TITLE: Rapid Learning or Feature Reuse? Towards Understanding the Effectiveness of MAML  
<https://openreview.net/forum?id=rkgMkCEtPB>  
AUTHORS: Aniruddh Raghu, Maithra Raghu, Samy Bengio, Oriol Vinyals  
HIGHLIGHT: The success of MAML relies on feature reuse from the meta-initialization, which also yields a natural simplification of the algorithm, with the inner loop removed for the network body, as well as other insights on the head and body.

338, TITLE: Augmenting Genetic Algorithms with Deep Neural Networks for Exploring the Chemical Space  
<https://openreview.net/forum?id=H1lmyRNFvr>  
AUTHORS: AkshatKumar Nigam, Pascal Friederich, Mario Krenn, Alan Aspuru-Guzik  
HIGHLIGHT: Tackling inverse design via genetic algorithms augmented with deep neural networks.

339, TITLE: Improved Sample Complexities for Deep Neural Networks and Robust Classification via an All-Layer Margin  
[https://openreview.net/forum?id=HJe\\_yR4Fwr](https://openreview.net/forum?id=HJe_yR4Fwr)  
AUTHORS: Colin Wei, Tengyu Ma  
HIGHLIGHT: We propose a new notion of margin that has a direct relationship with neural net generalization, and obtain improved generalization bounds for neural nets and robust classification by analyzing this margin.

340, TITLE: Identity Crisis: Memorization and Generalization Under Extreme Overparameterization  
<https://openreview.net/forum?id=B1l6y0VFPr>  
AUTHORS: Chiyuan Zhang, Samy Bengio, Moritz Hardt, Michael C. Mozer, Yoram Singer  
HIGHLIGHT: We propose a new notion of margin that has a direct relationship with neural net generalization, and obtain improved generalization bounds for neural nets and robust classification by analyzing this margin.

341, TITLE: ReMixMatch: Semi-Supervised Learning with Distribution Matching and Augmentation Anchoring  
<https://openreview.net/forum?id=Hk1keR4KPB>  
AUTHORS: David Berthelot, Nicholas Carlini, Ekin D. Cubuk, Alex Kurakin, Kihyuk Sohn, Han Zhang, Colin Raffel  
HIGHLIGHT: We introduce Distribution Matching and Augmentation Anchoring, two improvements to MixMatch which produce state-of-the-art results and enable surprisingly strong performance with only 40 labels on CIFAR-10 and SVHN.

342, TITLE: Adaptive Structural Fingerprints for Graph Attention Networks  
<https://openreview.net/forum?id=BJxWx0NYPr>  
AUTHORS: Kai Zhang, Yaokang Zhu, Jun Wang, Jie Zhang  
HIGHLIGHT: Exploiting rich structural details in graph-structured data via adaptive "structural fingerprints"

343, TITLE: CAQL: Continuous Action Q-Learning  
<https://openreview.net/forum?id=BkxXe0Etwr>  
AUTHORS: Moonkyung Ryu, Yinlam Chow, Ross Anderson, Christian Tjandraatmadja, Craig Boutilier  
HIGHLIGHT: A general framework of value-based reinforcement learning for continuous control

344, TITLE: Learning Heuristics for Quantified Boolean Formulas through Reinforcement Learning  
<https://openreview.net/forum?id=BJluxREKDB>  
AUTHORS: Gil Lederman, Markus Rabe, Sanjit Seshia, Edward A. Lee  
HIGHLIGHT: We use RL to automatically learn branching heuristic within a state of the art QBF solver, on industrial problems.

345, TITLE: Pure and Spurious Critical Points: a Geometric Study of Linear Networks  
<https://openreview.net/forum?id=rkgOICVYvB>  
AUTHORS: Matthew Trager, Kathl?n Kohn, Joan Bruna  
HIGHLIGHT: We use RL to automatically learn branching heuristic within a state of the art QBF solver, on industrial problems.

346, TITLE: Neural Text Generation With Unlikelihood Training  
<https://openreview.net/forum?id=SJeYe0NtvH>  
AUTHORS: Sean Welleck, Ilya Kulikov, Stephen Roller, Emily Dinan, Kyunghyun Cho, Jason Weston  
HIGHLIGHT: We use RL to automatically learn branching heuristic within a state of the art QBF solver, on industrial problems.

- 347, TITLE: Semi-Supervised Generative Modeling for Controllable Speech Synthesis  
<https://openreview.net/forum?id=rJeqeCEtvH>  
AUTHORS: Raza Habib, Soroosh Mariooryad, Matt Shannon, Eric Battenberg, RJ Ryan, Daisy Stanton, David Kao, Tom Bagby  
HIGHLIGHT: We use RL to automatically learn branching heuristic within a state of the art QBF solver, on industrial problems.
- 348, TITLE: Dynamic Time Lag Regression: Predicting What & When  
<https://openreview.net/forum?id=SkxybANtDB>  
AUTHORS: Mandar Chandorkar, Cyril Furtlehner, Bala Poduval, Enrico Camporeale, Michele Sebag  
HIGHLIGHT: We propose a new regression framework for temporal phenomena having non-stationary time-lag dependencies.
- 349, TITLE: Scalable Model Compression by Entropy Penalized Reparameterization  
<https://openreview.net/forum?id=HkxwW0EYDS>  
AUTHORS: Deniz Oktay, Johannes Ball?, Abhinav Shrivastava, Saurabh Singh  
HIGHLIGHT: An end-to-end trainable model compression method optimizing accuracy jointly with the expected model size.
- 350, TITLE: AMRL: Aggregated Memory For Reinforcement Learning  
<https://openreview.net/forum?id=Bk17bREtDr>  
AUTHORS: Jacob Beck, Kamil Ciosek, Sam Devlin, Sebastian Tschiatschek, Cheng Zhang, Katja Hofmann  
HIGHLIGHT: In Deep RL, order-invariant functions can be used in conjunction with standard memory modules to improve gradient decay and resilience to noise.
- 351, TITLE: Efficient Riemannian Optimization on the Stiefel Manifold via the Cayley Transform  
<https://openreview.net/forum?id=HJxV-ANKDH>  
AUTHORS: Jun Li, Fuxin Li, Sinisa Todorovic  
HIGHLIGHT: In Deep RL, order-invariant functions can be used in conjunction with standard memory modules to improve gradient decay and resilience to noise.
- 352, TITLE: UNPAIRED POINT CLOUD COMPLETION ON REAL SCANS USING ADVERSARIAL TRAINING  
<https://openreview.net/forum?id=HkgrZ0EYwB>  
AUTHORS: Xuelin Chen, Baoquan Chen, Niloy J. Mitra  
HIGHLIGHT: In Deep RL, order-invariant functions can be used in conjunction with standard memory modules to improve gradient decay and resilience to noise.
- 353, TITLE: Adjustable Real-time Style Transfer  
[https://openreview.net/forum?id=HJe\\_Z04Yvr](https://openreview.net/forum?id=HJe_Z04Yvr)  
AUTHORS: Mohammad Babaeizadeh, Golnaz Ghiasi  
HIGHLIGHT: Stochastic style transfer with adjustable features.
- 354, TITLE: Stochastic Weight Averaging in Parallel: Large-Batch Training That Generalizes Well  
<https://openreview.net/forum?id=rygFWAEfWS>  
AUTHORS: Vipul Gupta, Santiago Akle Serrano, Dennis DeCoste  
HIGHLIGHT: We propose SWAP, a distributed algorithm for large-batch training of neural networks.
- 355, TITLE: Short and Sparse Deconvolution --- A Geometric Approach  
<https://openreview.net/forum?id=Byg5ZANtvH>  
AUTHORS: Yenson Lau, Qing Qu, Han-Wen Kuo, Pengcheng Zhou, Yuqian Zhang, John Wright  
HIGHLIGHT: We propose SWAP, a distributed algorithm for large-batch training of neural networks.
- 356, TITLE: Selection via Proxy: Efficient Data Selection for Deep Learning  
<https://openreview.net/forum?id=HJg2b0VYDr>  
AUTHORS: Cody Coleman, Christopher Yeh, Stephen Mussmann, Baharan Mirzasoleiman, Peter Bailis, Percy Liang, Jure Leskovec, Matei Zaharia  
HIGHLIGHT: we can significantly improve the computational efficiency of data selection in deep learning by using a much smaller proxy model to perform data selection.
- 357, TITLE: Global Relational Models of Source Code

- <https://openreview.net/forum?id=B1InbRNtwr>  
AUTHORS: Vincent J. Hellendoorn, Charles Sutton, Rishabh Singh, Petros Maniatis  
HIGHLIGHT: Models of source code that combine global and structural features learn more powerful representations of programs.
- 358, TITLE: Detecting Extrapolation with Local Ensembles  
<https://openreview.net/forum?id=BJl6bANtwH>  
AUTHORS: David Madras, James Atwood, Alexander D'Amour  
HIGHLIGHT: We present local ensembles, a method for detecting extrapolation in trained models, which approximates the variance of an ensemble using local-second order information.
- 359, TITLE: Learning to Link  
<https://openreview.net/forum?id=S1eRbANtDB>  
AUTHORS: Maria-Florina Balcan, Travis Dick, Manuel Lang  
HIGHLIGHT: We show how to use data to automatically learn low-loss linkage procedures and metrics for specific clustering applications.
- 360, TITLE: Adversarially robust transfer learning  
<https://openreview.net/forum?id=ryebG04YvB>  
AUTHORS: Ali Shafahi, Parsa Saadatpanah, Chen Zhu, Amin Ghiasi, Christoph Studer, David Jacobs, Tom Goldstein  
HIGHLIGHT: Robust models have robust feature extractors which can be useful for transferring robustness to other domains
- 361, TITLE: Overlearning Reveals Sensitive Attributes  
<https://openreview.net/forum?id=SJeNz04tDS>  
AUTHORS: Congzheng Song, Vitaly Shmatikov  
HIGHLIGHT: Overlearning means that a model trained for a seemingly simple objective implicitly learns to recognize attributes and concepts that are (1) not part of the learning objective, and (2) sensitive from a privacy or bias perspective.
- 362, TITLE: Bridging Mode Connectivity in Loss Landscapes and Adversarial Robustness  
<https://openreview.net/forum?id=SJgwzCEKwH>  
AUTHORS: Pu Zhao, Pin-Yu Chen, Payel Das, Karthikeyan Natesan Ramamurthy, Xue Lin  
HIGHLIGHT: A novel approach using mode connectivity in loss landscapes to mitigate adversarial effects, repair tampered models and evaluate adversarial robustness
- 363, TITLE: Differentially Private Meta-Learning  
<https://openreview.net/forum?id=rJgqMRVYvr>  
AUTHORS: Jeffrey Li, Mikhail Khodak, Sebastian Caldas, Ameet Talwalkar  
HIGHLIGHT: A novel approach using mode connectivity in loss landscapes to mitigate adversarial effects, repair tampered models and evaluate adversarial robustness
- 364, TITLE: One-Shot Pruning of Recurrent Neural Networks by Jacobian Spectrum Evaluation  
<https://openreview.net/forum?id=r1e9GCNKvH>  
AUTHORS: Shunshi Zhang, Bradley C. Stadie  
HIGHLIGHT: New Objective for One-Shot Pruning Recurrent Neural Networks
- 365, TITLE: Meta-Dataset: A Dataset of Datasets for Learning to Learn from Few Examples  
<https://openreview.net/forum?id=rkgAGAVKPr>  
AUTHORS: Eleni Triantafillou, Tyler Zhu, Vincent Dumoulin, Pascal Lamblin, Utku Evci, Kelvin Xu, Ross Goroshin, Carles Gelada, Kevin Swersky, Pierre-Antoine Manzagol, Hugo Larochelle  
HIGHLIGHT: We propose a new large-scale diverse environment for few-shot learning, and evaluate popular models' performance on it, revealing important research challenges.
- 366, TITLE: Are Transformers universal approximators of sequence-to-sequence functions?  
<https://openreview.net/forum?id=ByxRM0Ntvr>  
AUTHORS: Chulhee Yun, Srinadh Bhojanapalli, Ankit Singh Rawat, Sashank Reddi, Sanjiv Kumar  
HIGHLIGHT: We prove that Transformer networks are universal approximators of sequence-to-sequence functions.
- 367, TITLE: GOING BEYOND TOKEN-LEVEL PRE-TRAINING FOR EMBEDDING-BASED LARGE-SCALE RETRIEVAL  
<https://openreview.net/forum?id=rkg-mA4FDr>

AUTHORS: Wei-Cheng Chang, Felix X. Yu, Yin-Wen Chang, Yiming Yang, Sanjiv Kumar  
HIGHLIGHT: We consider large-scale retrieval problems such as question answering retrieval and present a comprehensive study of how different sentence level pre-training improving the BERT-style token-level pre-training for two-tower Transformer models.

368, TITLE: Deep Imitative Models for Flexible Inference, Planning, and Control  
<https://openreview.net/forum?id=SkI4mRNYDr>  
AUTHORS: Nicholas Rhinehart, Rowan McAllister, Sergey Levine  
HIGHLIGHT: In this paper, we propose Imitative Models to combine the benefits of IL and goal-directed planning: probabilistic predictive models of desirable behavior able to plan interpretable expert-like trajectories to achieve specified goals.

369, TITLE: CM3: Cooperative Multi-goal Multi-stage Multi-agent Reinforcement Learning  
<https://openreview.net/forum?id=S11EX04tPr>  
AUTHORS: Jiachen Yang, Alireza Nakhaei, David Isele, Kikuo Fujimura, Hongyuan Zha  
HIGHLIGHT: A modular method for fully cooperative multi-goal multi-agent reinforcement learning, based on curriculum learning for efficient exploration and credit assignment for action-goal interactions.

370, TITLE: Robust And Interpretable Blind Image Denoising Via Bias-Free Convolutional Neural Networks  
<https://openreview.net/forum?id=HJISmC4FPS>  
AUTHORS: Sreyas Mohan, Zahra Kadkhodaie, Eero P. Simoncelli, Carlos Fernandez-Granda  
HIGHLIGHT: A modular method for fully cooperative multi-goal multi-agent reinforcement learning, based on curriculum learning for efficient exploration and credit assignment for action-goal interactions.

371, TITLE: Towards Better Understanding of Adaptive Gradient Algorithms in Generative Adversarial Nets  
<https://openreview.net/forum?id=SJxIm0VtwH>  
AUTHORS: Mingrui Liu, Youssef Mroueh, Jerret Ross, Wei Zhang, Xiaodong Cui, Payel Das, Tianbao Yang  
HIGHLIGHT: This paper provides novel analysis of adaptive gradient algorithms for solving non-convex non-concave min-max problems as GANs, and explains the reason why adaptive gradient methods outperform its non-adaptive counterparts by empirical studies.

372, TITLE: DeepV2D: Video to Depth with Differentiable Structure from Motion  
<https://openreview.net/forum?id=HJeO7RNKPr>  
AUTHORS: Zachary Teed, Jia Deng  
HIGHLIGHT: DeepV2D predicts depth from a video clip by composing elements of classical SfM into a fully differentiable network.

373, TITLE: Learning Space Partitions for Nearest Neighbor Search  
<https://openreview.net/forum?id=rkenmREFDr>  
AUTHORS: Yihe Dong, Piotr Indyk, Ilya Razenshteyn, Tal Wagner  
HIGHLIGHT: We use supervised learning (and in particular deep learning) to produce better space partitions for fast nearest neighbor search.

374, TITLE: Playing the lottery with rewards and multiple languages: lottery tickets in RL and NLP  
<https://openreview.net/forum?id=S1xnXRVFwH>  
AUTHORS: Haonan Yu, Sergey Edunov, Yuandong Tian, Ari S. Morcos  
HIGHLIGHT: We find that the lottery ticket phenomenon is present in both NLP and RL, and find that it can be used to train compressed Transformers to high performance

375, TITLE: Sign-OPT: A Query-Efficient Hard-label Adversarial Attack  
<https://openreview.net/forum?id=SkITQCNTvS>  
AUTHORS: Minhao Cheng, Simranjit Singh, Patrick H. Chen, Pin-Yu Chen, Sijia Liu, Cho-Jui Hsieh  
HIGHLIGHT: We find that the lottery ticket phenomenon is present in both NLP and RL, and find that it can be used to train compressed Transformers to high performance

376, TITLE: Toward Amortized Ranking-Critical Training For Collaborative Filtering  
<https://openreview.net/forum?id=HJxR7R4FvS>  
AUTHORS: Sam Lobel, Chunyuan Li, Jianfeng Gao, Lawrence Carin  
HIGHLIGHT: We apply the actor-critic methodology from reinforcement learning to collaborative filtering, resulting in improved performance across a variety of latent-variable models

- 377, TITLE: Intrinsic Motivation for Encouraging Synergistic Behavior  
<https://openreview.net/forum?id=SJleNCNtDH>  
AUTHORS: Rohan Chitnis, Shubham Tulsiani, Saurabh Gupta, Abhinav Gupta  
HIGHLIGHT: We propose a formulation of intrinsic motivation that is suitable as an exploration bias in multi-agent sparse-reward synergistic tasks, by encouraging agents to affect the world in ways that would not be achieved if they were acting individually.
- 378, TITLE: Chameleon: Adaptive Code Optimization For Expedited Deep Neural Network Compilation  
<https://openreview.net/forum?id=rygG4AVFvH>  
AUTHORS: Byung Hoon Ahn, Pranroy Pilligundla, Hadi Esmailzadeh  
HIGHLIGHT: Reinforcement learning and Adaptive Sampling for Optimized Compilation of Deep Neural Networks.
- 379, TITLE: The function of contextual illusions  
<https://openreview.net/forum?id=H1gB4RVKvB>  
AUTHORS: Drew Linsley, Junkyung Kim, Alekh Ashok, Thomas Serre  
HIGHLIGHT: Contextual illusions are a feature, not a bug, of neural routines optimized for contour detection.
- 380, TITLE: Locality and Compositionality in Zero-Shot Learning  
[https://openreview.net/forum?id=Hye\\_VONKwr](https://openreview.net/forum?id=Hye_VONKwr)  
AUTHORS: Tristan Sylvain, Linda Petrini, Devon Hjelm  
HIGHLIGHT: An analysis of the effects of compositionality and locality on representation learning for zero-shot learning.
- 381, TITLE: Understanding Knowledge Distillation in Non-autoregressive Machine Translation  
<https://openreview.net/forum?id=BygFVAEKDH>  
AUTHORS: Chunting Zhou, Jiatao Gu, Graham Neubig  
HIGHLIGHT: We systematically examine why knowledge distillation is crucial to the training of non-autoregressive translation (NAT) models, and propose methods to further improve the distilled data to best match the capacity of an NAT model.
- 382, TITLE: Thieves on Sesame Street! Model Extraction of BERT-based APIs  
<https://openreview.net/forum?id=Byl5NREFDr>  
AUTHORS: Kalpesh Krishna, Gaurav Singh Tomar, Ankur P. Parikh, Nicolas Papernot, Mohit Iyyer  
HIGHLIGHT: Outputs of modern NLP APIs on nonsensical text provide strong signals about model internals, allowing adversaries to steal the APIs.
- 383, TITLE: Fast is better than free: Revisiting adversarial training  
<https://openreview.net/forum?id=BJx040EFvH>  
AUTHORS: Eric Wong, Leslie Rice, J. Zico Kolter  
HIGHLIGHT: FGSM-based adversarial training, with randomization, works just as well as PGD-based adversarial training: we can use this to train a robust classifier in 6 minutes on CIFAR10, and 12 hours on ImageNet, on a single machine.
- 384, TITLE: DBA: Distributed Backdoor Attacks against Federated Learning  
<https://openreview.net/forum?id=rkgyS0VFvr>  
AUTHORS: Chulin Xie, Keli Huang, Pin-Yu Chen, Bo Li  
HIGHLIGHT: We proposed a novel distributed backdoor attack on federated learning and show that it is not only more effective compared with standard centralized attacks, but also harder to be defended by existing robust FL methods
- 385, TITLE: DeFINE: Deep Factorized Input Word Embeddings for Neural Sequence Modeling  
<https://openreview.net/forum?id=rJeXS04FPH>  
AUTHORS: Sachin Mehta, Rik Koncel-Kedziorski, Mohammad Rastegari, Hannaneh Hajishirzi  
HIGHLIGHT: DeFINE uses a deep, hierarchical, sparse network with new skip connections to learn better word embeddings efficiently.
- 386, TITLE: Sampling-Free Learning of Bayesian Quantized Neural Networks  
<https://openreview.net/forum?id=rylVHR4FPB>  
AUTHORS: Jiahao Su, Milan Cvitkovic, Furong Huang  
HIGHLIGHT: We propose Bayesian quantized networks, for which we learn a posterior distribution over their quantized parameters.
- 387, TITLE: Learning to solve the credit assignment problem  
<https://openreview.net/forum?id=ByeUBANtvB>

- AUTHORS: Benjamin James Lansdell, Prashanth Ravi Prakash, Konrad Paul Kording  
HIGHLIGHT: Perturbations can be used to train feedback weights to learn in fully connected and convolutional neural networks
- 388, TITLE: Four Things Everyone Should Know to Improve Batch Normalization  
<https://openreview.net/forum?id=HJx8HANFDH>  
AUTHORS: Cecilia Summers, Michael J. Dinneen  
HIGHLIGHT: Four things that improve batch normalization across all batch sizes
- 389, TITLE: Pseudo-LiDAR++: Accurate Depth for 3D Object Detection in Autonomous Driving  
<https://openreview.net/forum?id=BJedHRVtPB>  
AUTHORS: Yurong You\*, Yan Wang\*, Wei-Lun Chao\*, Divyansh Garg, Geoff Pleiss, Bharath Hariharan, Mark Campbell, Kilian Q. Weinberger  
HIGHLIGHT: Four things that improve batch normalization across all batch sizes
- 390, TITLE: SloMo: Improving Communication-Efficient Distributed SGD with Slow Momentum  
<https://openreview.net/forum?id=SkxJ8REYYPH>  
AUTHORS: Jianyu Wang, Vinayak Tantia, Nicolas Ballas, Michael Rabbat  
HIGHLIGHT: SlowMo improves the optimization and generalization performance of communication-efficient decentralized algorithms without sacrificing speed.
- 391, TITLE: MetaPix: Few-Shot Video Retargeting  
<https://openreview.net/forum?id=SJx1URNKwH>  
AUTHORS: Jessica Lee, Deva Ramanan, Rohit Girdhar  
HIGHLIGHT: Video retargeting typically requires large amount of target data to be effective, which may not always be available; we propose a metalearning approach that improves over popular baselines while producing temporally coherent frames.
- 392, TITLE: Learning to Learn by Zeroth-Order Oracle  
<https://openreview.net/forum?id=ryxz8CVYDH>  
AUTHORS: Yangjun Ruan, Yuanhao Xiong, Sashank Reddi, Sanjiv Kumar, Cho-Jui Hsieh  
HIGHLIGHT: Novel variant of learning to learn framework for zeroth-order optimization that learns both the update rule and the Gaussian sampling rule.
- 393, TITLE: Decentralized Distributed PPO: Mastering PointGoal Navigation  
<https://openreview.net/forum?id=H1gX8C4YPr>  
AUTHORS: Erik Wijmans, Abhishek Kadian, Ari Morcos, Stefan Lee, Irfan Essa, Devi Parikh, Manolis Savva, Dhruv Batra  
HIGHLIGHT: Novel variant of learning to learn framework for zeroth-order optimization that learns both the update rule and the Gaussian sampling rule.
- 394, TITLE: PAC Confidence Sets for Deep Neural Networks via Calibrated Prediction  
<https://openreview.net/forum?id=BJxVI04YvB>  
AUTHORS: Sangdon Park, Osbert Bastani, Nikolai Matni, Insup Lee  
HIGHLIGHT: Novel variant of learning to learn framework for zeroth-order optimization that learns both the update rule and the Gaussian sampling rule.
- 395, TITLE: Precision Gating: Improving Neural Network Efficiency with Dynamic Dual-Precision Activations  
<https://openreview.net/forum?id=SJgVU0EKwS>  
AUTHORS: Yichi Zhang, Ritchie Zhao, Weizhe Hua, Nayun Xu, Edward Suh, Zhiru Zhang  
HIGHLIGHT: We propose precision gating, an end-to-end trainable dual-precision activation quantization technique for deep neural networks.
- 396, TITLE: Locally Constant Networks  
<https://openreview.net/forum?id=Bke8UR4FPB>  
AUTHORS: Guang-He Lee, Tommi S. Jaakkola  
HIGHLIGHT: A novel neural architecture which implicitly learns an (oblique) decision tree.
- 397, TITLE: Span Recovery for Deep Neural Networks with Applications to Input Obfuscation  
<https://openreview.net/forum?id=B1guLAVFDB>  
AUTHORS: Rajesh Jayaram, David P. Woodruff, Qiuyi Zhang

**HIGHLIGHT:** We provably recover the span of a deep multi-layered neural network with latent structure and empirically apply efficient span recovery algorithms to attack networks by obfuscating inputs.

398, **TITLE:** Improving Neural Language Generation with Spectrum Control  
<https://openreview.net/forum?id=ByxY8CNtvr>  
**AUTHORS:** Lingxiao Wang, Jing Huang, Kevin Huang, Ziniu Hu, Guangtao Wang, Quanquan Gu  
**HIGHLIGHT:** We provably recover the span of a deep multi-layered neural network with latent structure and empirically apply efficient span recovery algorithms to attack networks by obfuscating inputs.

399, **TITLE:** Learn to Explain Efficiently via Neural Logic Inductive Learning  
<https://openreview.net/forum?id=SJlh8CEYDB>  
**AUTHORS:** Yuan Yang, Le Song  
**HIGHLIGHT:** An efficient differentiable ILP model that learns first-order logic rules that can explain the data.

400, **TITLE:** Improved memory in recurrent neural networks with sequential non-normal dynamics  
<https://openreview.net/forum?id=ryx1wRNFvB>  
**AUTHORS:** Emin Orhan, Xaq Pitkow  
**HIGHLIGHT:** a feedforward, chain-like motif (1-&gt;2-&gt;3-&gt;...) is proposed as a useful inductive bias for better memory in RNNs; amazingly, it works.

401, **TITLE:** Neural Module Networks for Reasoning over Text  
<https://openreview.net/forum?id=SygWvAVFPr>  
**AUTHORS:** Nitish Gupta, Kevin Lin, Dan Roth, Sameer Singh, Matt Gardner  
**HIGHLIGHT:** This paper extends neural module networks to answer compositional questions against text by introducing differentiable modules that perform reasoning over text and symbols in a probabilistic manner.

402, **TITLE:** Higher-Order Function Networks for Learning Composable 3D Object Representations  
<https://openreview.net/forum?id=HJgfDREKDB>  
**AUTHORS:** Eric Mitchell, Selim Engin, Volkan Isler, Daniel D Lee  
**HIGHLIGHT:** Neural nets can encode complex 3D objects into the parameters of other (surprisingly small) neural nets

403, **TITLE:** Variational Hetero-Encoder Randomized GANs for Joint Image-Text Modeling  
<https://openreview.net/forum?id=H1x5wRVtvS>  
**AUTHORS:** Hao Zhang, Bo Chen, Long Tian, Zhengjue Wang, Mingyuan Zhou  
**HIGHLIGHT:** A novel Bayesian deep learning framework that captures and relates hierarchical semantic and visual concepts, performing well on a variety of image and text modeling and generation tasks.

404, **TITLE:** Towards Fast Adaptation of Neural Architectures with Meta Learning  
<https://openreview.net/forum?id=r1eowANFvr>  
**AUTHORS:** Dongze Lian, Yin Zheng, Yintao Xu, Yanxiong Lu, Leyu Lin, Peilin Zhao, Junzhou Huang, Shenghua Gao  
**HIGHLIGHT:** A novel Bayesian deep learning framework that captures and relates hierarchical semantic and visual concepts, performing well on a variety of image and text modeling and generation tasks.

405, **TITLE:** Graph Constrained Reinforcement Learning for Natural Language Action Spaces  
<https://openreview.net/forum?id=B1x6w0EtwH>  
**AUTHORS:** Prithviraj Ammanabrolu, Matthew Hausknecht  
**HIGHLIGHT:** We present KG-A2C, a reinforcement learning agent that builds a dynamic knowledge graph while exploring and generates natural language using a template-based action space - outperforming all current agents on a wide set of text-based games.

406, **TITLE:** Prediction, Consistency, Curvature: Representation Learning for Locally-Linear Control  
[https://openreview.net/forum?id=BJxG\\_0EtDS](https://openreview.net/forum?id=BJxG_0EtDS)  
**AUTHORS:** Nir Levine, Yinlam Chow, Rui Shu, Ang Li, Mohammad Ghavamzadeh, Hung Bui  
**HIGHLIGHT:** Learning embedding for control with high-dimensional observations

407, **TITLE:** Augmenting Non-Collaborative Dialog Systems with Explicit Semantic and Strategic Dialog History  
<https://openreview.net/forum?id=ryxQuANKPB>  
**AUTHORS:** Yiheng Zhou, Yulia Tsvetkov, Alan W Black, Zhou Yu  
**HIGHLIGHT:** Learning embedding for control with high-dimensional observations



408, TITLE: BERTScore: Evaluating Text Generation with BERT  
<https://openreview.net/forum?id=SkeHuCVFDr>  
AUTHORS: Tianyi Zhang\*, Varsha Kishore\*, Felix Wu\*, Kilian Q. Weinberger, Yoav Artzi  
HIGHLIGHT: We propose BERTScore, an automatic evaluation metric for text generation, which correlates better with human judgments and provides stronger model selection performance than existing metrics.

409, TITLE: Neural Execution of Graph Algorithms  
<https://openreview.net/forum?id=SkgKOOEtvS>  
AUTHORS: Petar Velickovic, Rex Ying, Matilde Padovano, Raia Hadsell, Charles Blundell  
HIGHLIGHT: We supervise graph neural networks to imitate intermediate and step-wise outputs of classical graph algorithms, recovering highly favourable insights.

410, TITLE: On Need for Topology-Aware Generative Models for Manifold-Based Defenses  
[https://openreview.net/forum?id=r1IF\\_CeYwS](https://openreview.net/forum?id=r1IF_CeYwS)  
AUTHORS: Uyeong Jang, Susmit Jha, Somesh Jha  
HIGHLIGHT: We supervise graph neural networks to imitate intermediate and step-wise outputs of classical graph algorithms, recovering highly favourable insights.

411, TITLE: FSNet: Compression of Deep Convolutional Neural Networks by Filter Summary  
<https://openreview.net/forum?id=S1xtORNfWH>  
AUTHORS: Yingzhen Yang, Jiahui Yu, Nebojsa Jojic, Jun Huan, Thomas S. Huang  
HIGHLIGHT: We present a novel method of compression of deep Convolutional Neural Networks (CNNs) by weight sharing through a new representation of convolutional filters.

412, TITLE: Capsules with Inverted Dot-Product Attention Routing  
<https://openreview.net/forum?id=HJe6uANtwH>  
AUTHORS: Yao-Hung Hubert Tsai, Nitish Srivastava, Hanlin Goh, Ruslan Salakhutdinov  
HIGHLIGHT: We present a new routing method for Capsule networks, and it performs at-par with ResNet-18 on CIFAR-10/ CIFAR-100.

413, TITLE: Composition-based Multi-Relational Graph Convolutional Networks  
[https://openreview.net/forum?id=BylA\\_C4tPr](https://openreview.net/forum?id=BylA_C4tPr)  
AUTHORS: Shikhar Vashishth, Soumya Sanyal, Vikram Nitin, Partha Talukdar  
HIGHLIGHT: A Composition-based Graph Convolutional framework for multi-relational graphs.

414, TITLE: Gradient-Based Neural DAG Learning  
<https://openreview.net/forum?id=rklbKA4YDS>  
AUTHORS: S?bastien Lachapelle, Philippe Brouillard, Tristan Deleu, Simon Lacoste-Julien  
HIGHLIGHT: We are proposing a new score-based approach to structure/causal learning leveraging neural networks and a recent continuous constrained formulation to this problem

415, TITLE: The Local Elasticity of Neural Networks  
<https://openreview.net/forum?id=HJxMYANtPH>  
AUTHORS: Hangfeng He, Weijie Su  
HIGHLIGHT: We are proposing a new score-based approach to structure/causal learning leveraging neural networks and a recent continuous constrained formulation to this problem

416, TITLE: Composing Task-Agnostic Policies with Deep Reinforcement Learning  
<https://openreview.net/forum?id=H1ezFREtwH>  
AUTHORS: Ahmed H. Qureshi, Jacob J. Johnson, Yuzhe Qin, Taylor Henderson, Byron Boots, Michael C. Yip  
HIGHLIGHT: We propose a novel reinforcement learning-based skill transfer and composition method that takes the agent's primitive policies to solve unseen tasks.

417, TITLE: Convergence Behaviour of Some Gradient-Based Methods on Bilinear Zero-Sum Games  
<https://openreview.net/forum?id=SJIVY04FwH>  
AUTHORS: Guojun Zhang, Yaoliang Yu  
HIGHLIGHT: We systematically analyze the convergence behaviour of popular gradient algorithms for solving bilinear games, with both simultaneous and alternating updates.

- 418, TITLE: Discovering Motor Programs by Recomposing Demonstrations  
<https://openreview.net/forum?id=rkgHY0NYwr>  
AUTHORS: Tanmay Shankar, Shubham Tulsiani, Lerrel Pinto, Abhinav Gupta  
HIGHLIGHT: We learn a space of motor primitives from unannotated robot demonstrations, and show these primitives are semantically meaningful and can be composed for new robot tasks.
- 419, TITLE: Learning from Explanations with Neural Module Execution Tree  
<https://openreview.net/forum?id=rJIUt0EYwS>  
AUTHORS: Yujia Qin, Ziqi Wang, Wenxuan Zhou, Jun Yan, Qinyuan Ye, Xiang Ren, Leonardo Neves, Zhiyuan Liu  
HIGHLIGHT: We learn a space of motor primitives from unannotated robot demonstrations, and show these primitives are semantically meaningful and can be composed for new robot tasks.
- 420, TITLE: Jelly Bean World: A Testbed for Never-Ending Learning  
[https://openreview.net/forum?id=Byx\\_YAVYPH](https://openreview.net/forum?id=Byx_YAVYPH)  
AUTHORS: Emmanouil Antonios Platanios, Abulhair Saparov, Tom Mitchell  
HIGHLIGHT: We learn a space of motor primitives from unannotated robot demonstrations, and show these primitives are semantically meaningful and can be composed for new robot tasks.
- 421, TITLE: Coherent Gradients: An Approach to Understanding Generalization in Gradient Descent-based Optimization  
<https://openreview.net/forum?id=ryeFY0EFwS>  
AUTHORS: Sat Chatterjee  
HIGHLIGHT: We propose a hypothesis for why gradient descent generalizes based on how per-example gradients interact with each other.
- 422, TITLE: Probabilistic Connection Importance Inference and Lossless Compression of Deep Neural Networks  
<https://openreview.net/forum?id=HJgCF0Vfwr>  
AUTHORS: Xin Xing, Long Sha, Pengyu Hong, Zuofeng Shang, Jun S. Liu  
HIGHLIGHT: We propose a hypothesis for why gradient descent generalizes based on how per-example gradients interact with each other.
- 423, TITLE: MEMO: A Deep Network for Flexible Combination of Episodic Memories  
<https://openreview.net/forum?id=rJxl0EtDr>  
AUTHORS: Andrea Banino, Adrià Puigdomènech Badia, Raphael Koster, Martin J. Chadwick, Vinicius Zambaldi, Demis Hassabis, Caswell Barry, Matthew Botvinick, Dharshan Kumaran, Charles Blundell  
HIGHLIGHT: A memory architecture that support inferential reasoning.
- 424, TITLE: Economy Statistical Recurrent Units For Inferring Nonlinear Granger Causality  
<https://openreview.net/forum?id=SyxV9ANFDH>  
AUTHORS: Saurabh Khanna, Vincent Y. F. Tan  
HIGHLIGHT: A new recurrent neural network architecture for detecting pairwise Granger causality between nonlinearly interacting time series.
- 425, TITLE: Bayesian Meta Sampling for Fast Uncertainty Adaptation  
<https://openreview.net/forum?id=Bkxv90EKPB>  
AUTHORS: Zhenyi Wang, Yang Zhao, Ping Yu, Ruiyi Zhang, Changyou Chen  
HIGHLIGHT: We proposed a Bayesian meta sampling method for adapting the model uncertainty in meta learning
- 426, TITLE: Non-Autoregressive Dialog State Tracking  
[https://openreview.net/forum?id=H1e\\_cC4twS](https://openreview.net/forum?id=H1e_cC4twS)  
AUTHORS: Hung Le, Steven C.H. Hoi, Richard Socher  
HIGHLIGHT: We propose the first non-autoregressive neural model for Dialogue State Tracking (DST), achieving the SOTA accuracy (49.04%) on MultiWOZ2.1 benchmark, and reducing inference latency by an order of magnitude.
- 427, TITLE: Extreme Tensoring for Low-Memory Preconditioning  
<https://openreview.net/forum?id=SkIKcRNYDH>  
AUTHORS: Xinyi Chen, Naman Agarwal, Elad Hazan, Cyril Zhang, Yi Zhang  
HIGHLIGHT: We propose the first non-autoregressive neural model for Dialogue State Tracking (DST), achieving the SOTA accuracy (49.04%) on MultiWOZ2.1 benchmark, and reducing inference latency by an order of magnitude.
- 428, TITLE: Incremental RNN: A Dynamical View.

<https://openreview.net/forum?id=HylpqA4FwS>  
AUTHORS: Anil Kag, Ziming Zhang, Venkatesh Saligrama  
HIGHLIGHT: Incremental-RNNs resolves exploding/vanishing gradient problem by updating state vectors based on difference between previous state and that predicted by an ODE.

429, TITLE: The Early Phase of Neural Network Training  
<https://openreview.net/forum?id=Hk1liRNFwS>  
AUTHORS: Jonathan Frankle, David J. Schwab, Ari S. Morcos  
HIGHLIGHT: We thoroughly investigate neural network learning dynamics over the early phase of training, finding that these changes are crucial and difficult to approximate, though extended pretraining can recover them.

430, TITLE: NeurQuRI: Neural Question Requirement Inspector for Answerability Prediction in Machine Reading Comprehension  
<https://openreview.net/forum?id=ryxgsCVYPr>  
AUTHORS: Seohyun Back, Sai Chetan Chinthakindi, Akhil Kedia, Haejun Lee, Jaegul Choo  
HIGHLIGHT: We propose a neural question requirement inspection model called NeurQuRI that extracts a list of conditions from the question, each of which should be satisfied by the candidate answer generated by an MRC model.

431, TITLE: TOWARDS STABILIZING BATCH STATISTICS IN BACKWARD PROPAGATION OF BATCH NORMALIZATION  
<https://openreview.net/forum?id=SkGjRVKDS>  
AUTHORS: Junjie Yan, Ruosi Wan, Xiangyu Zhang, Wei Zhang, Yichen Wei, Jian Sun  
HIGHLIGHT: We propose a novel normalization method to handle small batch size cases.

432, TITLE: Single episode transfer for differing environmental dynamics in reinforcement learning  
<https://openreview.net/forum?id=rJeQoCNYDS>  
AUTHORS: Jiachen Yang, Brenden Petersen, Hongyuan Zha, Daniel Faissol  
HIGHLIGHT: Single episode policy transfer in a family of environments with related dynamics, via optimized probing for rapid inference of latent variables and immediate execution of a universal policy.

433, TITLE: Generalization through Memorization: Nearest Neighbor Language Models  
<https://openreview.net/forum?id=HkIBjCEKvH>  
AUTHORS: Urvashi Khadewal, Omer Levy, Dan Jurafsky, Luke Zettlemoyer, Mike Lewis  
HIGHLIGHT: We extend a pre-trained neural language model by linearly interpolating it with a k-nearest neighbors model, achieving new state-of-the-art results on Wikitext-103 with no additional training.

434, TITLE: Transformer-XH: Multi-hop question answering with eXtra Hop attention  
<https://openreview.net/forum?id=r1eliCNYwS>  
AUTHORS: Chen Zhao, Chenyan Xiong, Corby Rosset, Xia Song, Paul Bennett, Saurabh Tiwary  
HIGHLIGHT: We present Transformer-XH, which upgrades Transformer with eXtra Hop attentions to intrinsically model structured texts in a data driven way. It leads to a simpler yet state-of-the-art multi-hop QA system.

435, TITLE: Synthesizing Programmatic Policies that Inductively Generalize  
<https://openreview.net/forum?id=S118oANFDH>  
AUTHORS: Jeevana Priya Inala, Osbert Bastani, Zenna Tavares, Armando Solar-Lezama  
HIGHLIGHT: An approach to learn program policies that inductively generalize.

436, TITLE: Decoding As Dynamic Programming For Recurrent Autoregressive Models  
<https://openreview.net/forum?id=HklOo0VFDH>  
AUTHORS: Najam Zaidi, Trevor Cohn, Gholamreza Haffari  
HIGHLIGHT: Approximate inference using dynamic programming for Autoregressive models.

437, TITLE: Deep Double Descent: Where Bigger Models and More Data Hurt  
<https://openreview.net/forum?id=B1g5sA4twr>  
AUTHORS: Preetum Nakkiran, Gal Kaplun, Yamini Bansal, Tristan Yang, Boaz Barak, Ilya Sutskever  
HIGHLIGHT: We demonstrate, and characterize, realistic settings where bigger models are worse, and more data hurts.

438, TITLE: Intriguing Properties of Adversarial Training at Scale  
<https://openreview.net/forum?id=HyxJhCEFDS>  
AUTHORS: Cihang Xie, Alan Yuille

- HIGHLIGHT: The first rigor diagnose of large-scale adversarial training on ImageNet
- 439, TITLE: Shifted and Squeezed 8-bit Floating Point format for Low-Precision Training of Deep Neural Networks  
<https://openreview.net/forum?id=Bkxe2AVtPS>  
AUTHORS: Leopold Cambier, Anahita Bhiwandiwala, Ting Gong, Oguz H. Elibol, Mehran Nekuii, Hanlin Tang  
HIGHLIGHT: We propose a novel 8-bit format that eliminates the need for loss scaling, stochastic rounding, and other low precision techniques
- 440, TITLE: Distributed Bandit Learning: Near-Optimal Regret with Efficient Communication  
<https://openreview.net/forum?id=SJxZnR4YvB>  
AUTHORS: Yuanhao Wang, Jiachen Hu, Xiaoyu Chen, Liwei Wang  
HIGHLIGHT: We propose a novel 8-bit format that eliminates the need for loss scaling, stochastic rounding, and other low precision techniques
- 441, TITLE: Biologically inspired sleep algorithm for increased generalization and adversarial robustness in deep neural networks  
<https://openreview.net/forum?id=r1xGnA4Kvr>  
AUTHORS: Timothy Tadros, Giri Krishnan, Ramyaa Ramyaa, Maxim Bazhenov  
HIGHLIGHT: We describe a biologically inspired sleep algorithm for increased an artificial neural network's ability to extract the gist of a training set and exhibit increased robustness to adversarial attacks and general distortions.
- 442, TITLE: A Closer Look at the Optimization Landscapes of Generative Adversarial Networks  
<https://openreview.net/forum?id=HJeVnCEKwH>  
AUTHORS: Hugo Berard, Gauthier Gidel, Amjad Almahairi, Pascal Vincent, Simon Lacoste-Julien  
HIGHLIGHT: By proposing new visualization techniques we give better insights on GANs optimization in practical settings, we show that GANs on challenging datasets exhibit rotational behavior and do not converge to Nash-Equilibria
- 443, TITLE: On the Global Convergence of Training Deep Linear ResNets  
<https://openreview.net/forum?id=HJxEhREKDH>  
AUTHORS: Difan Zou, Philip M. Long, Quanquan Gu  
HIGHLIGHT: Under certain condition on the input and output linear transformations, both GD and SGD can achieve global convergence for training deep linear ResNets.
- 444, TITLE: Towards a Deep Network Architecture for Structured Smoothness  
<https://openreview.net/forum?id=Hklr204Fvr>  
AUTHORS: Haroun Habeeb, Oluwasanmi Koyejo  
HIGHLIGHT: A feedforward layer to incorporate structured smoothness into a deep learning model
- 445, TITLE: Revisiting Self-Training for Neural Sequence Generation  
<https://openreview.net/forum?id=SJgdnAVKDH>  
AUTHORS: Junxian He, Jiatao Gu, Jiajun Shen, Marc'Aurelio Ranzato  
HIGHLIGHT: We revisit self-training as a semi-supervised learning method for neural sequence generation problem, and show that self-training can be quite successful with injected noise.
- 446, TITLE: Denoising and Regularization via Exploiting the Structural Bias of Convolutional Generators  
<https://openreview.net/forum?id=HJeqhA4YDS>  
AUTHORS: Reinhard Heckel and Mahdi Soltanolkotabi  
HIGHLIGHT: We revisit self-training as a semi-supervised learning method for neural sequence generation problem, and show that self-training can be quite successful with injected noise.
- 447, TITLE: Variational Autoencoders for Highly Multivariate Spatial Point Processes Intensities  
<https://openreview.net/forum?id=B1lj20NFDS>  
AUTHORS: Baichuan Yuan, Xiaowei Wang, Andrea Bertozzi, Hongxia Yang  
HIGHLIGHT: We revisit self-training as a semi-supervised learning method for neural sequence generation problem, and show that self-training can be quite successful with injected noise.
- 448, TITLE: Model-Augmented Actor-Critic: Backpropagating through Paths  
<https://openreview.net/forum?id=Skln2A4YDB>  
AUTHORS: Ignasi Clavera, Yao Fu, Pieter Abbeel

**HIGHLIGHT:** Policy gradient through backpropagation through time using learned models and Q-functions. SOTA results in reinforcement learning benchmark environments.

449, **TITLE:** LambdaNet: Probabilistic Type Inference using Graph Neural Networks

<https://openreview.net/forum?id=Hkx6hANtwH>

**AUTHORS:** Jiayi Wei, Maruth Goyal, Greg Durrett, Isil Dillig

**HIGHLIGHT:** Policy gradient through backpropagation through time using learned models and Q-functions. SOTA results in reinforcement learning benchmark environments.

450, **TITLE:** From Inference to Generation: End-to-end Fully Self-supervised Generation of Human Face from Speech

<https://openreview.net/forum?id=H1guaREYPr>

**AUTHORS:** Hyeong-Seok Choi, Changdae Park, Kyogu Lee

**HIGHLIGHT:** This paper proposes a method of end-to-end multi-modal generation of human face from speech based on a self-supervised learning framework.

451, **TITLE:** Visual Representation Learning with 3D View-Contrastive Inverse Graphics Networks

<https://openreview.net/forum?id=BJxt60VtPr>

**AUTHORS:** Adam W. Harley, Fangyu Li, Shrinidhi K. Lakshmikanth, Xian Zhou, Hsiao-Yu Fish Tung, Katerina

Fragkiadaki

**HIGHLIGHT:** We show that with the right loss and architecture, view-predictive learning improves 3D object detection

452, **TITLE:** Decoupling Representation and Classifier for Long-Tailed Recognition

<https://openreview.net/forum?id=r1gRTCvFvB>

**AUTHORS:** Bingyi Kang, Saining Xie, Marcus Rohrbach, Zhicheng Yan, Albert Gordo, Jiashi Feng, Yannis Kalantidis

**HIGHLIGHT:** We show that with the right loss and architecture, view-predictive learning improves 3D object detection

453, **TITLE:** Robust Reinforcement Learning for Continuous Control with Model Misspecification

<https://openreview.net/forum?id=HJgC60EtwB>

**AUTHORS:** Daniel J. Mankowitz, Nir Levine, Rae Jeong, Abbas Abdolmaleki, Jost Tobias Springenberg, Yuanyuan Shi,

Jackie Kay, Todd Hester, Timothy Mann, Martin Riedmiller

**HIGHLIGHT:** A framework for incorporating robustness to model misspecification into continuous control Reinforcement Learning algorithms.

454, **TITLE:** Cross-lingual Alignment vs Joint Training: A Comparative Study and A Simple Unified Framework

<https://openreview.net/forum?id=S11-C0NtwS>

**AUTHORS:** Zirui Wang\*, Jiateng Xie\*, Ruochen Xu, Yiming Yang, Graham Neubig, Jaime G. Carbonell

**HIGHLIGHT:** We conduct a comparative study of cross-lingual alignment vs joint training methods and unify these two previously exclusive paradigms in a new framework.

455, **TITLE:** Training Recurrent Neural Networks Online by Learning Explicit State Variables

<https://openreview.net/forum?id=SJgmR0NKPr>

**AUTHORS:** Somjit Nath, Vincent Liu, Alan Chan, Adam White, Martha White

**HIGHLIGHT:** We conduct a comparative study of cross-lingual alignment vs joint training methods and unify these two previously exclusive paradigms in a new framework.

456, **TITLE:** Uncertainty-guided Continual Learning with Bayesian Neural Networks

<https://openreview.net/forum?id=HklUCCVKDB>

**AUTHORS:** Sayna Ebrahimi, Mohamed Elhoseiny, Trevor Darrell, Marcus Rohrbach

**HIGHLIGHT:** A regularization-based approach for continual learning using Bayesian neural networks to predict parameters' importance

457, **TITLE:** Curriculum Loss: Robust Learning and Generalization against Label Corruption

<https://openreview.net/forum?id=rkgt0REKwS>

**AUTHORS:** Yueming Lyu, Ivor W. Tsang

**HIGHLIGHT:** A novel loss bridges curriculum learning and robust learning

458, **TITLE:** Picking Winning Tickets Before Training by Preserving Gradient Flow

<https://openreview.net/forum?id=SkgsACVKPH>

**AUTHORS:** Chaoqi Wang, Guodong Zhang, Roger Grosse

**HIGHLIGHT:** We introduced a pruning criterion for pruning networks before training by preserving gradient flow.

459, TITLE: Generative Models for Effective ML on Private, Decentralized Datasets  
<https://openreview.net/forum?id=SJgaRA4FPH>  
AUTHORS: Sean Augenstein, H. Brendan McMahan, Daniel Ramage, Swaroop Ramaswamy, Peter Kairouz, Mingqing Chen, Rajiv Mathews, Blaise Aguera y Arcas  
HIGHLIGHT: Generative Models + Federated Learning + Differential Privacy gives data scientists a way to analyze private, decentralized data (e.g., on mobile devices) where direct inspection is prohibited.

460, TITLE: Inductive representation learning on temporal graphs  
<https://openreview.net/forum?id=rJeWlyHYwH>  
AUTHORS: da Xu, chuanwei ruan, evren korpeoglu, sushant kumar, kannan acharya  
HIGHLIGHT: Generative Models + Federated Learning + Differential Privacy gives data scientists a way to analyze private, decentralized data (e.g., on mobile devices) where direct inspection is prohibited.

461, TITLE: BatchEnsemble: an Alternative Approach to Efficient Ensemble and Lifelong Learning  
<https://openreview.net/forum?id=SklflyrYDr>  
AUTHORS: Yeming Wen, Dustin Tran, Jimmy Ba  
HIGHLIGHT: We introduced BatchEnsemble, an efficient method for ensembling and lifelong learning which can be used to improve the accuracy and uncertainty of any neural network like typical ensemble methods.

462, TITLE: Towards neural networks that provably know when they don't know  
<https://openreview.net/forum?id=ByxGkySKwH>  
AUTHORS: Alexander Meinke, Matthias Hein  
HIGHLIGHT: We introduced BatchEnsemble, an efficient method for ensembling and lifelong learning which can be used to improve the accuracy and uncertainty of any neural network like typical ensemble methods.

463, TITLE: Iterative energy-based projection on a normal data manifold for anomaly localization  
<https://openreview.net/forum?id=HJx8lySKwr>  
AUTHORS: David Dehaene, Oriel Frigo, S?bastien Combrexelle, Pierre Eline  
HIGHLIGHT: We use gradient descent on a regularized autoencoder loss to correct anomalous images.

464, TITLE: Towards Stable and Efficient Training of Verifiably Robust Neural Networks  
<https://openreview.net/forum?id=Skxuk1rFwB>  
AUTHORS: Huan Zhang, Hongge Chen, Chaowei Xiao, Sven Gowal, Robert Stanforth, Bo Li, Duane Boning, Cho-Jui Hsieh  
HIGHLIGHT: We propose a new certified adversarial training method, CROWN-IBP, that achieves state-of-the-art robustness for  $L_\infty$  norm adversarial perturbations.

465, TITLE: Frequency-based Search-control in Dyna  
<https://openreview.net/forum?id=BlgskyStwr>  
AUTHORS: Yangchen Pan, Jincheng Mei, Amir-massoud Farahmand, Martha White  
HIGHLIGHT: Acquire states from high frequency region for search-control in Dyna.

466, TITLE: Learning representations for binary-classification without backpropagation  
<https://openreview.net/forum?id=Bke61krFvS>  
AUTHORS: Mathias Lechner  
HIGHLIGHT: First feedback alignment algorithm with provable learning guarantees for networks with single output neuron

467, TITLE: Polylogarithmic width suffices for gradient descent to achieve arbitrarily small test error with shallow ReLU networks  
<https://openreview.net/forum?id=HygegyrYwH>  
AUTHORS: Ziwei Ji, Matus Telgarsky  
HIGHLIGHT: First feedback alignment algorithm with provable learning guarantees for networks with single output neuron

468, TITLE: Physics-aware Difference Graph Networks for Sparsely-Observed Dynamics  
<https://openreview.net/forum?id=r1gelyrtwH>  
AUTHORS: Sungyong Seo\*, Chuizheng Meng\*, Yan Liu  
HIGHLIGHT: We propose physics-aware difference graph networks designed to effectively learn spatial differences to modeling sparsely-observed dynamics.

- 469, TITLE: HiLLOc: lossless image compression with hierarchical latent variable models  
<https://openreview.net/forum?id=r1IZgyBYwS>  
AUTHORS: James Townsend, Thomas Bird, Julius Kunze, David Barber  
HIGHLIGHT: We scale up lossless compression with latent variables, beating existing approaches on full-size ImageNet images.
- 470, TITLE: IMPACT: Importance Weighted Asynchronous Architectures with Clipped Target Networks  
<https://openreview.net/forum?id=BJeGJStPr>  
AUTHORS: Michael Luo, Jiahao Yao, Richard Liaw, Eric Liang, Ion Stoica  
HIGHLIGHT: IMPACT helps RL agents train faster by decreasing training wall-clock time and increasing sample efficiency simultaneously.
- 471, TITLE: On Bonus Based Exploration Methods In The Arcade Learning Environment  
<https://openreview.net/forum?id=BJewlyStDr>  
AUTHORS: Adrien Ali Taiga, William Fedus, Marlos C. Machado, Aaron Courville, Marc G. Bellemare  
HIGHLIGHT: We find that existing bonus-based exploration methods have not been able to address the exploration-exploitation trade-off in the Arcade Learning Environment.
- 472, TITLE: Adaptive Correlated Monte Carlo for Contextual Categorical Sequence Generation  
<https://openreview.net/forum?id=r1IOgyrKDS>  
AUTHORS: Xinjie Fan, Yizhe Zhang, Zhendong Wang, Mingyuan Zhou  
HIGHLIGHT: We find that existing bonus-based exploration methods have not been able to address the exploration-exploitation trade-off in the Arcade Learning Environment.
- 473, TITLE: Smoothness and Stability in GANs  
<https://openreview.net/forum?id=HJeOekHKwr>  
AUTHORS: Casey Chu, Kentaro Minami, Kenji Fukumizu  
HIGHLIGHT: We develop a principled theoretical framework for understanding and enforcing the stability of various types of GANs
- 474, TITLE: SNOW: Subscribing to Knowledge via Channel Pooling for Transfer & Lifelong Learning  
<https://openreview.net/forum?id=rJxtgJBKDr>  
AUTHORS: Chungkuk Yoo, Bumsoo Kang, Minsik Cho  
HIGHLIGHT: We propose SNOW, an efficient way of transfer and lifelong learning by subscribing knowledge of a source model for new tasks through a novel channel pooling block.
- 475, TITLE: Empirical Studies on the Properties of Linear Regions in Deep Neural Networks  
<https://openreview.net/forum?id=SkeF11HKwr>  
AUTHORS: Xiao Zhang, Dongrui Wu  
HIGHLIGHT: We propose SNOW, an efficient way of transfer and lifelong learning by subscribing knowledge of a source model for new tasks through a novel channel pooling block.
- 476, TITLE: Black-box Off-policy Estimation for Infinite-Horizon Reinforcement Learning  
<https://openreview.net/forum?id=S1ltg1rFDS>  
AUTHORS: Ali Mousavi, Lihong Li, Qiang Liu, Denny Zhou  
HIGHLIGHT: We present a novel approach for the off-policy estimation problem in infinite-horizon RL.
- 477, TITLE: PairNorm: Tackling Oversmoothing in GNNs  
<https://openreview.net/forum?id=rkecl1rtwB>  
AUTHORS: Lingxiao Zhao, Leman Akoglu  
HIGHLIGHT: We proposed a normalization layer for GNN models to solve the oversmoothing problem.
- 478, TITLE: Unsupervised Clustering using Pseudo-semi-supervised Learning  
<https://openreview.net/forum?id=rJlnxkSYPS>  
AUTHORS: Divam Gupta, Ramachandran Ramjee, Nipun Kwatra, Muthian Sivathanu  
HIGHLIGHT: Using ensembles and pseudo labels for unsupervised clustering
- 479, TITLE: Simple and Effective Regularization Methods for Training on Noisily Labeled Data with Generalization Guarantee

- <https://openreview.net/forum?id=Hke3gyHYwH>  
AUTHORS: Wei Hu, Zhiyuan Li, Dingli Yu  
HIGHLIGHT: Using ensembles and pseudo labels for unsupervised clustering
- 480, TITLE: Controlling generative models with continuous factors of variations  
<https://openreview.net/forum?id=H1laeJrKDB>  
AUTHORS: Antoine Plumerault, Hervé Le Borgne, Clément Hudelot  
HIGHLIGHT: A model to control the generation of images with GAN and beta-VAE with regard to scale and position of the objects
- 481, TITLE: Symplectic ODE-Net: Learning Hamiltonian Dynamics with Control  
<https://openreview.net/forum?id=ryxmb1rKDS>  
AUTHORS: Yaofeng Desmond Zhong, Biswadip Dey, Amit Chakraborty  
HIGHLIGHT: This work enforces Hamiltonian dynamics with control to learn system models from embedded position and velocity data, and exploits this physically-consistent dynamics to synthesize model-based control via energy shaping.
- 482, TITLE: Understanding l4-based Dictionary Learning: Interpretation, Stability, and Robustness  
<https://openreview.net/forum?id=SJeY-1BKDS>  
AUTHORS: Yuexiang Zhai, Hermish Mehta, Zhengyuan Zhou, Yi Ma  
HIGHLIGHT: We compare the l4-norm based dictionary learning with PCA, ICA and show its stability as well as robustness.
- 483, TITLE: Quantum Algorithms for Deep Convolutional Neural Networks  
<https://openreview.net/forum?id=Hygab1rKDS>  
AUTHORS: Iordanis Kerenidis, Jonas Landman, Anupam Prakash  
HIGHLIGHT: We provide the first algorithm for quantum computers implementing universal convolutional neural network with a speedup
- 484, TITLE: Self-Supervised Learning of Appliance Usage  
<https://openreview.net/forum?id=B1lJzyStvS>  
AUTHORS: Chen-Yu Hsu, Abbas Zeitoun, Guang-He Lee, Dina Katabi, Tommi Jaakkola  
HIGHLIGHT: We learn appliance usage patterns in homes without labels, using self-supervised learning with energy and location data
- 485, TITLE: Deep Graph Matching Consensus  
<https://openreview.net/forum?id=HyeJf1HKvS>  
AUTHORS: Matthias Fey, Jan E. Lenssen, Christopher Morris, Jonathan Masci, Nils M. Kriege  
HIGHLIGHT: We develop a deep graph matching architecture which refines initial correspondences based on a neighborhood consensus error.
- 486, TITLE: Beyond Linearization: On Quadratic and Higher-Order Approximation of Wide Neural Networks  
<https://openreview.net/forum?id=rklIGyBFPH>  
AUTHORS: Yu Bai, Jason D. Lee  
HIGHLIGHT: Wide neural networks can escape the NTK regime and couple with quadratic models, with provably nice optimization landscape and better generalization.
- 487, TITLE: Dynamic Sparse Training: Find Efficient Sparse Network From Scratch With Trainable Masked Layers  
<https://openreview.net/forum?id=SJlbGJrtDB>  
AUTHORS: Junjie LIU, Zhe XU, Runbin SHI, Ray C. C. Cheung, Hayden K.H. So  
HIGHLIGHT: We present a novel network pruning method that can find the optimal sparse structure during the training process with trainable pruning threshold
- 488, TITLE: Triple Wins: Boosting Accuracy, Robustness and Efficiency Together by Enabling Input-Adaptive Inference  
<https://openreview.net/forum?id=rJgzzJHtDB>  
AUTHORS: Ting-Kuei Hu, Tianlong Chen, Haotao Wang, Zhangyang Wang  
HIGHLIGHT: Is it possible to co-design model accuracy, robustness and efficiency to achieve their triple wins? Yes!
- 489, TITLE: Neural Policy Gradient Methods: Global Optimality and Rates of Convergence  
<https://openreview.net/forum?id=BJgQfkSYDS>  
AUTHORS: Lingxiao Wang, Qi Cai, Zhuoran Yang, Zhaoran Wang  
HIGHLIGHT: Is it possible to co-design model accuracy, robustness and efficiency to achieve their triple wins? Yes!



- 490, TITLE: Double Neural Counterfactual Regret Minimization  
<https://openreview.net/forum?id=ByedzkrKvH>  
AUTHORS: Hui Li, Kailiang Hu, Shaohua Zhang, Yuan Qi, Le Song  
HIGHLIGHT: We proposed a double neural framework to solve large-scale imperfect information game.
- 491, TITLE: GraphAF: a Flow-based Autoregressive Model for Molecular Graph Generation  
<https://openreview.net/forum?id=S1esMkHYPr>  
AUTHORS: Chence Shi\*, Minkai Xu\*, Zhaocheng Zhu, Weinan Zhang, Ming Zhang, Jian Tang  
HIGHLIGHT: A flow-based autoregressive model for molecular graph generation. Reaching state-of-the-art results on molecule generation and properties optimization.
- 492, TITLE: The Gambler's Problem and Beyond  
<https://openreview.net/forum?id=HyxnMyBKwB>  
AUTHORS: Baoxiang Wang, Shuai Li, Jiajin Li, Siu On Chan  
HIGHLIGHT: The optimal value function is fractal and is like a Cantor function.
- 493, TITLE: Multilingual Alignment of Contextual Word Representations  
<https://openreview.net/forum?id=r1xCMyBtPS>  
AUTHORS: Steven Cao, Nikita Kitaev, Dan Klein  
HIGHLIGHT: We propose procedures for evaluating and strengthening contextual embedding alignment and show that they both improve multilingual BERT's zero-shot XNLI transfer and provide useful insights into the model.
- 494, TITLE: The Curious Case of Neural Text Degeneration  
<https://openreview.net/forum?id=rygGQyrFvH>  
AUTHORS: Ari Holtzman, Jan Buys, Leo Du, Maxwell Forbes, Yejin Choi  
HIGHLIGHT: Current language generation systems either aim for high likelihood and devolve into generic repetition or miscalibrate their stochasticity?we provide evidence of both and propose a solution: Nucleus Sampling.
- 495, TITLE: Graph Convolutional Reinforcement Learning  
<https://openreview.net/forum?id=HkxdQkSYDB>  
AUTHORS: Jiechuan Jiang, Chen Dun, Tiejun Huang, Zongqing Lu  
HIGHLIGHT: Current language generation systems either aim for high likelihood and devolve into generic repetition or miscalibrate their stochasticity?we provide evidence of both and propose a solution: Nucleus Sampling.
- 496, TITLE: Meta-Learning Deep Energy-Based Memory Models  
<https://openreview.net/forum?id=SyljQyBFDH>  
AUTHORS: Sergey Bartunov, Jack Rae, Simon Osindero, Timothy Lillicrap  
HIGHLIGHT: Deep associative memory models using arbitrary neural networks as a storage.
- 497, TITLE: Exploratory Not Explanatory: Counterfactual Analysis of Saliency Maps for Deep RL  
<https://openreview.net/forum?id=rkl3m1BFDB>  
AUTHORS: Akanksha Atrey, Kaleigh Clary, David Jensen  
HIGHLIGHT: Proposing a new counterfactual-based methodology to evaluate the hypotheses generated from saliency maps about deep RL agent behavior.
- 498, TITLE: Fast Neural Network Adaptation via Parameters Remapping  
<https://openreview.net/forum?id=rklTmyBKPH>  
AUTHORS: Jiemin Fang\*, Yuzhu Sun\*, Kangjian Peng\*, Qian Zhang, Yuan Li, Wenyu Liu, Xinggang Wang  
HIGHLIGHT: Proposing a new counterfactual-based methodology to evaluate the hypotheses generated from saliency maps about deep RL agent behavior.
- 499, TITLE: Guiding Program Synthesis by Learning to Generate Examples  
<https://openreview.net/forum?id=BJl07ySKvS>  
AUTHORS: Larissa Laich, Pavol Bielik, Martin Vechev  
HIGHLIGHT: Proposing a new counterfactual-based methodology to evaluate the hypotheses generated from saliency maps about deep RL agent behavior.
- 500, TITLE: SNODE: Spectral Discretization of Neural ODEs for System Identification

- <https://openreview.net/forum?id=Sye0XkBKvS>  
AUTHORS: Alessio Quaglino, Marco Gallieri, Jonathan Masci, Jan Koutn?k  
HIGHLIGHT: This paper proposes the use of spectral element methods for fast and accurate training of Neural Ordinary Differential Equations for system identification.
- 501, TITLE: Generalized Convolutional Forest Networks for Domain Generalization and Visual Recognition  
<https://openreview.net/forum?id=H1lxVyStPH>  
AUTHORS: Jongbin Ryu, GiTaek Kwon, Ming-Hsuan Yang, Jongwoo Lim  
HIGHLIGHT: This paper proposes the use of spectral element methods for fast and accurate training of Neural Ordinary Differential Equations for system identification.
- 502, TITLE: Once for All: Train One Network and Specialize it for Efficient Deployment  
<https://openreview.net/forum?id=HyIxElHKwS>  
AUTHORS: Han Cai, Chuang Gan, Tianzhe Wang, Zhekai Zhang, Song Han  
HIGHLIGHT: We introduce techniques to train a single once-for-all network that fits many hardware platforms.
- 503, TITLE: Multi-Agent Interactions Modeling with Correlated Policies  
<https://openreview.net/forum?id=B1gZV1HYvS>  
AUTHORS: Minghuan Liu, Ming Zhou, Weinan Zhang, Yuzheng Zhuang, Jun Wang, Wulong Liu, Yong Yu  
HIGHLIGHT: Modeling complex multi-agent interactions under multi-agent imitation learning framework with explicit modeling of correlated policies by approximating opponents' policies.
- 504, TITLE: PCMC-Net: Feature-based Pairwise Choice Markov Chains  
<https://openreview.net/forum?id=BJgWE1SFwS>  
AUTHORS: Alix Lh?ritier  
HIGHLIGHT: We propose a generic neural network architecture equipping Pairwise Choice Markov Chains choice models with amortized and automatic differentiation based inference using alternatives' and individuals' features.
- 505, TITLE: Implementing Inductive bias for different navigation tasks through diverse RNN attractors  
<https://openreview.net/forum?id=Byx4NkrtDS>  
AUTHORS: Tie XU, Omri Barak  
HIGHLIGHT: Task agnostic pre-training can shape RNN's attractor landscape, and form diverse inductive bias for different navigation tasks
- 506, TITLE: Query2box: Reasoning over Knowledge Graphs in Vector Space Using Box Embeddings  
<https://openreview.net/forum?id=BJgr4kSFDS>  
AUTHORS: Hongyu Ren\*, Weihua Hu\*, Jure Leskovec  
HIGHLIGHT: Answering a wide class of logical queries over knowledge graphs with box embeddings in vector space
- 507, TITLE: Rethinking the Hyperparameters for Fine-tuning  
<https://openreview.net/forum?id=B1g8VkhFPH>  
AUTHORS: Hao Li, Pratik Chaudhari, Hao Yang, Michael Lam, Avinash Ravichandran, Rahul Bhotik, Stefano Soatto  
HIGHLIGHT: This paper re-examines several common practices of setting hyper-parameters for fine-tuning.
- 508, TITLE: Plug and Play Language Model: A simple baseline for controlled language generation  
<https://openreview.net/forum?id=H1edEyBKDS>  
AUTHORS: Sumanth Dathathri, Andrea Madotto, Janice Lan, Jane Hung, Eric Frank, Piero Molino, Jason Yosinski, Rosanne Liu  
HIGHLIGHT: We control the topic and sentiment of text generation (almost) without any training.
- 509, TITLE: Provable Benefit of Orthogonal Initialization in Optimizing Deep Linear Networks  
<https://openreview.net/forum?id=rkgqN1SYvr>  
AUTHORS: Wei Hu, Lechao Xiao, Jeffrey Pennington  
HIGHLIGHT: We provide for the first time a rigorous proof that orthogonal initialization speeds up convergence relative to Gaussian initialization, for deep linear networks.
- 510, TITLE: RGBD-GAN: Unsupervised 3D Representation Learning From Natural Image Datasets via RGBD Image Synthesis  
<https://openreview.net/forum?id=HyxjNyrPr>  
AUTHORS: Atsuhiko Noguchi, Tatsuya Harada

HIGHLIGHT: RGBD image generation for unsupervised camera parameter conditioning

511, TITLE: Towards Verified Robustness under Text Deletion Interventions  
<https://openreview.net/forum?id=SyxhVkrYvr>  
AUTHORS: Johannes Welbl, Po-Sen Huang, Robert Stanforth, Sven Gowal, Krishnamurthy (Dj) Dvijotham, Martin Szummer, Pushmeet Kohli  
HIGHLIGHT: Formal verification of a specification on a model's prediction undersensitivity using Interval Bound Propagation

512, TITLE: Jacobian Adversarially Regularized Networks for Robustness  
<https://openreview.net/forum?id=Hke0V1rKPS>  
AUTHORS: Alvin Chan, Yi Tay, Yew Soon Ong, Jie Fu  
HIGHLIGHT: We show that training classifiers to produce salient input Jacobian matrices with a GAN-like regularization can boost adversarial robustness.

513, TITLE: Thinking While Moving: Deep Reinforcement Learning with Concurrent Control  
<https://openreview.net/forum?id=SJexHkSFPS>  
AUTHORS: Ted Xiao, Eric Jang, Dmitry Kalashnikov, Sergey Levine, Julian Ibarz, Karol Hausman, Alexander Herzog  
HIGHLIGHT: Reinforcement learning formulation that allows agents to think and act at the same time, demonstrated on real-world robotic grasping.

514, TITLE: Evolutionary Population Curriculum for Scaling Multi-Agent Reinforcement Learning  
<https://openreview.net/forum?id=SJxbHkrKDH>  
AUTHORS: Qian Long\*, Zihan Zhou\*, Abhinav Gupta, Fei Fang, Yi Wu?, Xiaolong Wang?  
HIGHLIGHT: Reinforcement learning formulation that allows agents to think and act at the same time, demonstrated on real-world robotic grasping.

515, TITLE: ELECTRA: Pre-training Text Encoders as Discriminators Rather Than Generators  
<https://openreview.net/forum?id=r1xMH1BtvB>  
AUTHORS: Kevin Clark, Minh-Thang Luong, Quoc V. Le, Christopher D. Manning  
HIGHLIGHT: A text encoder trained to distinguish real input tokens from plausible fakes efficiently learns effective language representations.

516, TITLE: Emergent Systematic Generalization In a Situated Agent  
<https://openreview.net/forum?id=SkIGryBtwr>  
AUTHORS: Felix Hill, Andrew Lampinen, Rosalia Schneider, Stephen Clark, Matthew Botvinick, James L. McClelland, Adam Santoro  
HIGHLIGHT: We isolate the environmental and training factors that contribute to strong emergent systematic generalization in a situated language-learning agent

517, TITLE: Abstract Diagrammatic Reasoning with Multiplex Graph Networks  
<https://openreview.net/forum?id=ByxQB1BKwH>  
AUTHORS: Duo Wang, Mateja Jamnik, Pietro Lio  
HIGHLIGHT: MXGNet is a multilayer, multiplex graph based architecture which achieves good performance on various diagrammatic reasoning tasks.

518, TITLE: A Baseline for Few-Shot Image Classification  
<https://openreview.net/forum?id=rylXBkrYDS>  
AUTHORS: Guneet Singh Dhillon, Pratik Chaudhari, Avinash Ravichandran, Stefano Soatto  
HIGHLIGHT: Transductive fine-tuning of a deep network is a strong baseline for few-shot image classification and outperforms the state-of-the-art on all standard benchmarks.

519, TITLE: Learning to Retrieve Reasoning Paths over Wikipedia Graph for Question Answering  
<https://openreview.net/forum?id=SJgVHkrYDH>  
AUTHORS: Akari Asai, Kazuma Hashimoto, Hannaneh Hajishirzi, Richard Socher, Caiming Xiong  
HIGHLIGHT: Graph-based recurrent retriever that learns to retrieve reasoning paths over Wikipedia Graph outperforms the most recent state of the art on HotpotQA by more than 10 points.

520, TITLE: Pad? Activation Units: End-to-end Learning of Flexible Activation Functions in Deep Networks  
<https://openreview.net/forum?id=BJIBSKHtDS>  
AUTHORS: Alejandro Molina, Patrick Schramowski, Kristian Kersting

**HIGHLIGHT:** We introduce PAU, a new learnable activation function for neural networks. They free the network designers from the activation selection process and increase the test prediction accuracy.

521, **TITLE:** A FRAMEWORK FOR ROBUSTNESS CERTIFICATION OF SMOOTHED CLASSIFIERS USING F-DIVERGENCES  
<https://openreview.net/forum?id=SJKrkSFPH>  
**AUTHORS:** Krishnamurthy (Dj) Dvijotham, Jamie Hayes, Borja Balle, Zico Kolter, Chongli Qin, Andras Gyorgy, Kai Xiao, Sven Gowal, Pushmeet Kohli  
**HIGHLIGHT:** Develop a general framework to establish certified robustness of ML models against various classes of adversarial perturbations

522, **TITLE:** Contrastive Representation Distillation  
<https://openreview.net/forum?id=SkgpBJrtvS>  
**AUTHORS:** Yonglong Tian, Dilip Krishnan, Phillip Isola  
**HIGHLIGHT:** Representation/knowledge distillation by maximizing mutual information between teacher and student

523, **TITLE:** Certified Defenses for Adversarial Patches  
<https://openreview.net/forum?id=HyeaSkYYPH>  
**AUTHORS:** Ping-yeh Chiang\*, Renkun Ni\*, Ahmed Abdelkader, Chen Zhu, Chris Studor, Tom Goldstein  
**HIGHLIGHT:** Representation/knowledge distillation by maximizing mutual information between teacher and student

524, **TITLE:** Sample Efficient Policy Gradient Methods with Recursive Variance Reduction  
<https://openreview.net/forum?id=HJlxIJBFDr>  
**AUTHORS:** Pan Xu, Felicia Gao, Quanquan Gu  
**HIGHLIGHT:** Representation/knowledge distillation by maximizing mutual information between teacher and student

525, **TITLE:** Deep Symbolic Superoptimization Without Human Knowledge  
<https://openreview.net/forum?id=r1egIyBFPS>  
**AUTHORS:** Hui Shi, Yang Zhang, Xinyun Chen, Yuandong Tian, Jishen Zhao  
**HIGHLIGHT:** Representation/knowledge distillation by maximizing mutual information between teacher and student

526, **TITLE:** Explain Your Move: Understanding Agent Actions Using Focused Feature Saliency  
<https://openreview.net/forum?id=SJgzLkKBPB>  
**AUTHORS:** Piyush Gupta, Nikaash Puri, Sukriti Verma, Dhruv Kayastha, Shripad Deshmukh, Balaji Krishnamurthy, Sameer Singh  
**HIGHLIGHT:** We propose a model-agnostic approach to explain the behaviour of black-box deep RL agents, trained to play Atari and board games, by highlighting relevant features of an input state.

527, **TITLE:** Universal Approximation with Certified Networks  
<https://openreview.net/forum?id=BlgX8kBTPr>  
**AUTHORS:** Maximilian Baader, Matthew Mirman, Martin Vechev  
**HIGHLIGHT:** We prove that for a large class of functions  $f$  there exists an interval certified robust network approximating  $f$  up to arbitrary precision.

528, **TITLE:** Measuring and Improving the Use of Graph Information in Graph Neural Networks  
<https://openreview.net/forum?id=rkelIkHKvS>  
**AUTHORS:** Yifan Hou, Jian Zhang, James Cheng, Kaili Ma, Richard T. B. Ma, Hongzhi Chen, Ming-Chang Yang  
**HIGHLIGHT:** We prove that for a large class of functions  $f$  there exists an interval certified robust network approximating  $f$  up to arbitrary precision.

529, **TITLE:** State-only Imitation with Transition Dynamics Mismatch  
<https://openreview.net/forum?id=HJgLLyrYwB>  
**AUTHORS:** Tanmay Gangwani, Jian Peng  
**HIGHLIGHT:** Algorithm for imitation with state-only expert demonstrations; builds on adversarial-IRL; experiments with transition dynamics mismatch b/w expert and imitator

530, **TITLE:** Adversarial AutoAugment  
<https://openreview.net/forum?id=ByxdUySKvS>  
**AUTHORS:** Xinyu Zhang, Qiang Wang, Jian Zhang, Zhao Zhong

HIGHLIGHT: We introduce the idea of adversarial learning into automatic data augmentation to improve the generalization of a target network.

531, TITLE: Meta Dropout: Learning to Perturb Latent Features for Generalization

<https://openreview.net/forum?id=BJgd81SYwr>

AUTHORS: Hae Beom Lee, Taewook Nam, Eunho Yang, Sung Ju Hwang

HIGHLIGHT: We introduce the idea of adversarial learning into automatic data augmentation to improve the generalization of a target network.

532, TITLE: R<sup>2</sup>-Fair Inference

<https://openreview.net/forum?id=HkgsUJrtDB>

AUTHORS: Sina Baharlouei, Maher Nouiehed, Meisam Razaviyayn

HIGHLIGHT: We introduce the idea of adversarial learning into automatic data augmentation to improve the generalization of a target network.

533, TITLE: Learning transport cost from subset correspondence

<https://openreview.net/forum?id=SJIRUkrFPS>

AUTHORS: Ruishan Liu, Akshay Balsubramani, James Zou

HIGHLIGHT: We introduce the idea of adversarial learning into automatic data augmentation to improve the generalization of a target network.

534, TITLE: BlockSwap: Fisher-guided Block Substitution for Network Compression on a Budget

<https://openreview.net/forum?id=SkldkSFPB>

AUTHORS: Jack Turner, Elliot J. Crowley, Michael O'Boyle, Amos Storkey, Gavin Gray

HIGHLIGHT: A simple and effective method for reducing large neural networks to flexible parameter targets based on block substitution.

535, TITLE: Variance Reduction With Sparse Gradients

<https://openreview.net/forum?id=Syx1DkSYwB>

AUTHORS: Melih Elibol, Lihua Lei, Michael I. Jordan

HIGHLIGHT: We use sparsity to improve the computational complexity of variance reduction methods.

536, TITLE: Abductive Commonsense Reasoning

<https://openreview.net/forum?id=Byg1v1HKDB>

AUTHORS: Chandra Bhagavatula, Ronan Le Bras, Chaitanya Malaviya, Keisuke Sakaguchi, Ari Holtzman, Hannah

Rashkin, Doug Downey, Wen-tau Yih, Yejin Choi

HIGHLIGHT: We use sparsity to improve the computational complexity of variance reduction methods.

537, TITLE: Discrepancy Ratio: Evaluating Model Performance When Even Experts Disagree on the Truth

<https://openreview.net/forum?id=Byg-wJSYDS>

AUTHORS: Igor Lovchinsky, Alon Daks, Israel Malkin, Pouya Samangouei, Ardavan Saeeedi, Yang Liu, Swami

Sankaranarayanan, Tomer Gafner, Ben Sternlieb, Patrick Maher, Nathan Silberman

HIGHLIGHT: A framework for evaluating model performance when even experts disagree on what the ground truth is.

538, TITLE: Weakly Supervised Disentanglement with Guarantees

<https://openreview.net/forum?id=HJgSwyBKvr>

AUTHORS: Rui Shu, Yining Chen, Abhishek Kumar, Stefano Ermon, Ben Poole

HIGHLIGHT: We construct a theoretical framework for weakly supervised disentanglement and conducted lots of experiments to back up the theory.

539, TITLE: Nesterov Accelerated Gradient and Scale Invariance for Adversarial Attacks

<https://openreview.net/forum?id=SJIHwkBYDH>

AUTHORS: Jiadong Lin, Chuanbiao Song, Kun He, Liwei Wang, John E. Hopcroft

HIGHLIGHT: We proposed a Nesterov Iterative Fast Gradient Sign Method (NI-FGSM) and a Scale-Invariant attack Method (SIM) that can boost the transferability of adversarial examples for image classification.

540, TITLE: Fantastic Generalization Measures and Where to Find Them

<https://openreview.net/forum?id=SJgIPJBFvH>

AUTHORS: Yiding Jiang\*, Behnam Neyshabur\*, Dilip Krishnan, Hossein Mobahi, Samy Bengio

**HIGHLIGHT:** We empirically study generalization measures over more than 2000 models, identify common pitfall in existing practice of studying generalization measures and provide some new bounds based on measures in our study.

541, **TITLE:** Robustness Verification for Transformers  
<https://openreview.net/forum?id=BJxwPJHFWs>  
**AUTHORS:** Zhouxing Shi, Huan Zhang, Kai-Wei Chang, Minlie Huang, Cho-Jui Hsieh  
**HIGHLIGHT:** We propose the first algorithm for verifying the robustness of Transformers.

542, **TITLE:** A Simple Randomization Technique for Generalization in Deep Reinforcement Learning  
<https://openreview.net/forum?id=HJgcvJBFvB>  
**AUTHORS:** Kimin Lee, Kibok Lee, Jinwoo Shin, Honglak Lee  
**HIGHLIGHT:** We propose a simple randomization technique for improving generalization in deep reinforcement learning across tasks with various unseen visual patterns.

543, **TITLE:** Tensor Decompositions for Temporal Knowledge Base Completion  
<https://openreview.net/forum?id=rke2P1BFwS>  
**AUTHORS:** Timothée Lacroix, Guillaume Obozinski, Nicolas Usunier  
**HIGHLIGHT:** We propose new tensor decompositions and associated regularizers to obtain state of the art performances on temporal knowledge base completion.

544, **TITLE:** On Universal Equivariant Set Networks  
<https://openreview.net/forum?id=HkxTwkrKDB>  
**AUTHORS:** Nimrod Segol, Yaron Lipman  
**HIGHLIGHT:** Settling permutation equivariance universality for popular deep models.

545, **TITLE:** Provable robustness against all adversarial  $l_p$ -perturbations for  $l_q$   
[https://openreview.net/forum?id=rklk\\_ySYPB](https://openreview.net/forum?id=rklk_ySYPB)  
**AUTHORS:** Francesco Croce, Matthias Hein  
**HIGHLIGHT:** We introduce a method to train models with provable robustness wrt all the  $l_p$ -norms for  $l_q$  simultaneously.

546, **TITLE:** Don't Use Large Mini-batches, Use Local SGD  
<https://openreview.net/forum?id=BlEyO1BFPr>  
**AUTHORS:** Tao Lin, Sebastian U. Stich, Kumar Kshitij Patel, Martin Jaggi  
**HIGHLIGHT:** We introduce a method to train models with provable robustness wrt all the  $l_p$ -norms for  $l_q$  simultaneously.

547, **TITLE:** Kernel of CycleGAN as a principal homogeneous space  
<https://openreview.net/forum?id=BlEWOJHKvB>  
**AUTHORS:** Nikita Moriakov, Jonas Adler, Jonas Teuwen  
**HIGHLIGHT:** The space of approximate solutions of CycleGAN admits a lot of symmetry, and an identity loss does not fix this.

548, **TITLE:** Distributionally Robust Neural Networks  
<https://openreview.net/forum?id=ryxGuJrFvS>  
**AUTHORS:** Shiori Sagawa\*, Pang Wei Koh\*, Tatsunori B. Hashimoto, Percy Liang  
**HIGHLIGHT:** Overparameterized neural networks can be distributionally robust, but only when you account for generalization.

549, **TITLE:** On Solving Minimax Optimization Locally: A Follow-the-Ridge Approach  
[https://openreview.net/forum?id=Hkx7\\_1rKwS](https://openreview.net/forum?id=Hkx7_1rKwS)  
**AUTHORS:** Yuanhao Wang, Guodong Zhang, Jimmy Ba  
**HIGHLIGHT:** Overparameterized neural networks can be distributionally robust, but only when you account for generalization.

550, **TITLE:** A Neural Dirichlet Process Mixture Model for Task-Free Continual Learning  
<https://openreview.net/forum?id=SJxSOJStPr>  
**AUTHORS:** Soochan Lee, Junsoo Ha, Dongsu Zhang, Gunhee Kim

**HIGHLIGHT:** We propose an expansion-based approach for task-free continual learning for the first time. Our model consists of a set of neural network experts and expands the number of experts under the Bayesian nonparametric principle.

551, **TITLE:** Hyper-SAGNN: a self-attention based graph neural network for hypergraphs  
<https://openreview.net/forum?id=ryeHuJBtPH>  
**AUTHORS:** Ruochi Zhang, Yuesong Zou, Jian Ma  
**HIGHLIGHT:** We develop a new self-attention based graph neural network called Hyper-SAGNN applicable to homogeneous and heterogeneous hypergraphs with variable hyperedge sizes that can fulfill tasks like node classification and hyperedge prediction.

552, **TITLE:** Neural Epitome Search for Architecture-Agnostic Network Compression  
<https://openreview.net/forum?id=HyxjOyrKvr>  
**AUTHORS:** Daquan Zhou, Xiaojie Jin, Qibin Hou, Kaixin Wang, Jianchao Yang, Jiashi Feng  
**HIGHLIGHT:** We present a novel neural network compression method which can reuse the parameters efficiently to reduce the model size.

553, **TITLE:** On the Equivalence between Node Embeddings and Structural Graph Representations  
<https://openreview.net/forum?id=SJxzFySKwH>  
**AUTHORS:** Balasubramaniam Srinivasan, Bruno Ribeiro  
**HIGHLIGHT:** We develop the foundations of a unifying theoretical framework connecting node embeddings and structural graph representations through invariant theory

554, **TITLE:** Probability Calibration for Knowledge Graph Embedding Models  
<https://openreview.net/forum?id=S1g8K1BFwS>  
**AUTHORS:** Pedro Tabacof, Luca Costabello  
**HIGHLIGHT:** We propose a novel method to calibrate knowledge graph embedding models without the need of negative examples.

555, **TITLE:** Why Not to Use Zero Imputation? Correcting Sparsity Bias in Training Neural Networks  
<https://openreview.net/forum?id=BylsKkHYvH>  
**AUTHORS:** Joonyoung Yi, Juhyuk Lee, Sung Ju Hwang, Eunho Yang  
**HIGHLIGHT:** We propose a novel method to calibrate knowledge graph embedding models without the need of negative examples.

556, **TITLE:** DropEdge: Towards Deep Graph Convolutional Networks on Node Classification  
<https://openreview.net/forum?id=Hkx1qkrKPr>  
**AUTHORS:** Yu Rong, Wenbing Huang, Tingyang Xu, Junzhou Huang  
**HIGHLIGHT:** This paper proposes DropEdge, a novel and flexible technique to alleviate over-smoothing and overfitting issue in deep Graph Convolutional Networks.

557, **TITLE:** Masked Based Unsupervised Content Transfer  
<https://openreview.net/forum?id=BJe-91BtvH>  
**AUTHORS:** Ron Mokady, Sagie Benaim, Lior Wolf, Amit Bermano  
**HIGHLIGHT:** This paper proposes DropEdge, a novel and flexible technique to alleviate over-smoothing and overfitting issue in deep Graph Convolutional Networks.

558, **TITLE:** U-GAT-IT: Unsupervised Generative Attentional Networks with Adaptive Layer-Instance Normalization for Image-to-Image Translation  
<https://openreview.net/forum?id=BJIZ5ySKPH>  
**AUTHORS:** Junho Kim, Minjae Kim, Hyeonwoo Kang, Kwang Hee Lee  
**HIGHLIGHT:** This paper proposes DropEdge, a novel and flexible technique to alleviate over-smoothing and overfitting issue in deep Graph Convolutional Networks.

559, **TITLE:** Inductive and Unsupervised Representation Learning on Graph Structured Objects  
<https://openreview.net/forum?id=rkem91rtDB>  
**AUTHORS:** Lichen Wang, Bo Zong, Qianqian Ma, Wei Cheng, Jingchao Ni, Wenchao Yu, Yanchi Liu, Dongjin Song, Haifeng Chen, Yun Fu  
**HIGHLIGHT:** This paper proposed a novel framework for graph similarity learning in inductive and unsupervised scenario.

560, **TITLE:** Batch-shaping for learning conditional channel gated networks  
<https://openreview.net/forum?id=Bke89JBtvB>

- AUTHORS: Babak Ehteshami Bejnordi, Tijmen Blankevoort, Max Welling  
HIGHLIGHT: A method that trains large capacity neural networks with significantly improved accuracy and lower dynamic computational cost
- 561, TITLE: Learning Robust Representations via Multi-View Information Bottleneck  
<https://openreview.net/forum?id=B1xwcyHFDr>  
AUTHORS: Marco Federici, Anjan Dutta, Patrick Forr?, Nate Kushman, Zeynep Akata  
HIGHLIGHT: We extend the information bottleneck method to the unsupervised multiview setting and show state of the art results on standard datasets
- 562, TITLE: Deep probabilistic subsampling for task-adaptive compressed sensing  
<https://openreview.net/forum?id=SJeq9JBFvH>  
AUTHORS: Iris A.M. Huijben, Bastiaan S. Veeling, Ruud J.G. van Sloun  
HIGHLIGHT: We extend the information bottleneck method to the unsupervised multiview setting and show state of the art results on standard datasets
- 563, TITLE: Robust anomaly detection and backdoor attack detection via differential privacy  
<https://openreview.net/forum?id=SJx0q1rtvS>  
AUTHORS: Min Du, Ruoxi Jia, Dawn Song  
HIGHLIGHT: This paper shows that differential privacy could improve the utility of outlier detection, novelty detection and backdoor attack detection, through both a theoretical analysis and extensive experimental results (constructed and real-world).
- 564, TITLE: Learning to Guide Random Search  
<https://openreview.net/forum?id=B1gHokBKwS>  
AUTHORS: Ozan Sener, Vladlen Koltun  
HIGHLIGHT: We improve the sample-efficiency of the random search for functions defined on low-dimensional manifolds. Our method jointly learns the underlying manifold and optimizes the function.
- 565, TITLE: Lagrangian Fluid Simulation with Continuous Convolutions  
<https://openreview.net/forum?id=B1IDoJSYDH>  
AUTHORS: Benjamin Ummenhofer, Lukas Prantl, Nils Th?rey, Vladlen Koltun  
HIGHLIGHT: We learn particle-based fluid simulation with convolutional networks.
- 566, TITLE: Reinforced Genetic Algorithm Learning for Optimizing Computation Graphs  
<https://openreview.net/forum?id=rkxDoJBYPB>  
AUTHORS: Aditya Paliwal, Felix Gimeno, Vinod Nair, Yujia Li, Miles Lubin, Pushmeet Kohli, Oriol Vinyals  
HIGHLIGHT: We use deep RL to learn a policy that directs the search of a genetic algorithm to better optimize the execution cost of computation graphs, and show improved results on real-world TensorFlow graphs.
- 567, TITLE: Compressive Transformers for Long-Range Sequence Modelling  
<https://openreview.net/forum?id=SylKikSYDH>  
AUTHORS: Jack W. Rae, Anna Potapenko, Siddhant M. Jayakumar, Chloe Hillier, Timothy P. Lillicrap  
HIGHLIGHT: Long-range transformer using a compressive memory, achieves sota in wikitext-103 and enwik8 LM benchmarks, release a new book-level LM benchmark PG-19.  
To promote the domain of long-range sequence learning, we propose a new open-vocabulary language modelling benchmark derived from books, PG-19.
- 568, TITLE: A Stochastic Derivative Free Optimization Method with Momentum  
<https://openreview.net/forum?id=HylAoJSKvH>  
AUTHORS: Eduard Gorbunov, Adel Bibi, Ozan Sener, El Houcine Bergou, Peter Richtarik  
HIGHLIGHT: We develop and analyze a new derivative free optimization algorithm with momentum and importance sampling with applications to continuous control.
- 569, TITLE: Understanding and Improving Information Transfer in Multi-Task Learning  
<https://openreview.net/forum?id=SylzhkbtDB>  
AUTHORS: Sen Wu, Hongyang Zhang, Christopher R?  
HIGHLIGHT: A Theoretical Study of Multi-Task Learning with Practical Implications for Improving Multi-Task Training and Transfer Learning
- 570, TITLE: Learning To Explore Using Active Neural Mapping



- <https://openreview.net/forum?id=HkIXnIBKDH>  
AUTHORS: Devendra Singh Chaplot, Saurabh Gupta, Dhiraj Gandhi, Abhinav Gupta, Ruslan Salakhutdinov  
HIGHLIGHT: A modular and hierarchical approach to learn policies for exploring 3D environments.
- 571, TITLE: EMPIR: Ensembles of Mixed Precision Deep Networks for Increased Robustness Against Adversarial Attacks  
<https://openreview.net/forum?id=HJem3yHKwH>  
AUTHORS: Sanchari Sen, Balaraman Ravindran, Anand Raghunathan  
HIGHLIGHT: We propose ensembles of mixed-precision DNNs as a new form of defense against adversarial attacks
- 572, TITLE: Quantifying Point-Prediction Uncertainty in Neural Networks via Residual Estimation with an I/O Kernel  
<https://openreview.net/forum?id=rkxNh1Stvr>  
AUTHORS: Xin Qiu, Elliot Meyerson, Risto Miikkilainen  
HIGHLIGHT: Learning to Estimate Point-Prediction Uncertainty and Correct Output in Neural Networks
- 573, TITLE: B-Spline CNNs on Lie groups  
<https://openreview.net/forum?id=H1gBhkBFDH>  
AUTHORS: Erik J Bekkers  
HIGHLIGHT: The paper describes a flexible framework for building CNNs that are equivariant to a large class of transformations groups.
- 574, TITLE: Neural Outlier Rejection for Self-Supervised Keypoint Learning  
<https://openreview.net/forum?id=Skx82ySYPH>  
AUTHORS: Jiexiong Tang, Rares Ambrus, Vitor Guizilini, Hanme Kim  
HIGHLIGHT: Learning to extract distinguishable keypoints from a proxy task, outlier rejection.
- 575, TITLE: Reducing Transformer Depth on Demand with Structured Dropout  
<https://openreview.net/forum?id=SyIO2yStDr>  
AUTHORS: Angela Fan, Edouard Grave, Armand Joulin  
HIGHLIGHT: Layerdrop, a form of structured dropout that allows you to train one model at training time and prune to any desired depth at test time. You can also
- 576, TITLE: Cross-Lingual Ability of Multilingual BERT: An Empirical Study  
<https://openreview.net/forum?id=HJeT3yrtDr>  
AUTHORS: Karthikeyan K, Zihan Wang, Stephen Mayhew, Dan Roth  
HIGHLIGHT: Cross-Lingual Ability of Multilingual BERT: An Empirical Study
- 577, TITLE: Spatially Parallel Attention and Component Extraction for Scene Decomposition  
<https://openreview.net/forum?id=rkl03ySYDH>  
AUTHORS: Sungjin Ahn, Zhixuan Lin, Weihao Sun, Skand Vishwanath Peri, Gautam Singh, Yi-Fu Wu, Fei Deng, Jindong Jiang  
HIGHLIGHT: We propose a generative latent variable model for unsupervised scene decomposition that provides factorized object representation per foreground object while also decomposing background segments of complex morphology.
- 578, TITLE: RIDE: Rewarding Impact-Driven Exploration for Procedurally-Generated Environments  
<https://openreview.net/forum?id=rkg-TJBFpB>  
AUTHORS: Roberta Raileanu, Tim Rocktäschel  
HIGHLIGHT: Instead of rewarding agents for predicting the next state, reward them for taking actions that lead to changes in the state.
- 579, TITLE: On the geometry and learning low-dimensional embeddings for directed graphs  
<https://openreview.net/forum?id=SkxQp1StDH>  
AUTHORS: Thorben Funke, Tian Guo, Alen Lancic, Nino Antulov-Fantulin  
HIGHLIGHT: We propose a novel node embedding of directed graphs to statistical manifolds and analyze connections to divergence, geometry and efficient learning procedure.
- 580, TITLE: Efficient Probabilistic Logic Reasoning with Graph Neural Networks  
<https://openreview.net/forum?id=rJg76kStwH>  
AUTHORS: Yuyu Zhang, Xinshi Chen, Yuan Yang, Arun Ramamurthy, Bo Li, Yuan Qi, Le Song  
HIGHLIGHT: We employ graph neural networks in the variational EM framework for efficient inference and learning of Markov Logic Networks.

- 581, TITLE: GraphSAINT: Graph Sampling Based Inductive Learning Method  
<https://openreview.net/forum?id=BJe8pkHFwS>  
AUTHORS: Hanqing Zeng, Hongkuan Zhou, Ajitesh Srivastava, Rajgopal Kannan, Viktor Prasanna  
HIGHLIGHT: We propose a graph sampling based minibatch construction method for training Graph Convolutional Networks.
- 582, TITLE: You Only Train Once: Loss-Conditional Training of Deep Networks  
<https://openreview.net/forum?id=HyxY6JHKwr>  
AUTHORS: Alexey Dosovitskiy, Josip Djolonga  
HIGHLIGHT: A method to train a single model simultaneously minimizing a family of loss functions instead of training a set of per-loss models.
- 583, TITLE: Projection Based Constrained Policy Optimization  
<https://openreview.net/forum?id=rke3TJrtPS>  
AUTHORS: Tsung-Yen Yang, Justinian Rosca, Karthik Narasimhan, Peter J. Ramadge  
HIGHLIGHT: We propose a new algorithm that learns constraint-satisfying policies, and provide theoretical analysis and empirical demonstration in the context of reinforcement learning with constraints.
- 584, TITLE: Infinite-Horizon Differentiable Model Predictive Control  
<https://openreview.net/forum?id=ryxC6kSYPr>  
AUTHORS: Sebastian East, Marco Gallieri, Jonathan Masci, Jan Koutnik, Mark Cannon  
HIGHLIGHT: We propose a new algorithm that learns constraint-satisfying policies, and provide theoretical analysis and empirical demonstration in the context of reinforcement learning with constraints.
- 585, TITLE: Combining Q-Learning and Search with Amortized Value Estimates  
<https://openreview.net/forum?id=SkeAaJrKDS>  
AUTHORS: Jessica B. Hamrick, Victor Bapst, Alvaro Sanchez-Gonzalez, Tobias Pfaff, Theophane Weber, Lars Buesing, Peter W. Battaglia  
HIGHLIGHT: We propose a model-based method called "Search with Amortized Value Estimates" (SAVE) which leverages both real and planned experience by combining Q-learning with Monte-Carlo Tree Search, achieving strong performance with very small search budgets.
- 586, TITLE: Training Generative Adversarial Networks from Incomplete Observations using Factorised Discriminators  
<https://openreview.net/forum?id=Hye1RJHKwB>  
AUTHORS: Daniel Stoller, Sebastian Ewert, Simon Dixon  
HIGHLIGHT: We decompose the discriminator in a GAN in a principled way so that each component can be independently trained on different parts of the input. The resulting "FactorGAN" can be used for semi-supervised learning and in missing data scenarios.
- 587, TITLE: Decentralized Deep Learning with Arbitrary Communication Compression  
<https://openreview.net/forum?id=SkgGckrKvH>  
AUTHORS: Anastasia Koloskova\*, Tao Lin\*, Sebastian U Stich, Martin Jaggi  
HIGHLIGHT: We propose Choco-SGD---decentralized SGD with compressed communication---for non-convex objectives and show its strong performance in various deep learning applications (on-device learning, datacenter case).
- 588, TITLE: Toward Evaluating Robustness of Deep Reinforcement Learning with Continuous Control  
<https://openreview.net/forum?id=SyIL0krYPS>  
AUTHORS: Tsui-Wei Weng, Krishnamurthy (Dj) Dvijotham\*, Jonathan Uesato\*, Kai Xiao\*, Sven Gowal\*, Robert Stanforth\*, Pushmeet Kohli  
HIGHLIGHT: We study the problem of continuous control agents in deep RL with adversarial attacks and proposed a two-step algorithm based on learned model dynamics.
- 589, TITLE: Gradient  $\ell_1$  Regularization for Quantization Robustness  
<https://openreview.net/forum?id=ryxK0JBtPr>  
AUTHORS: Milad Alizadeh, Arash Behboodi, Mart van Baalen, Christos Louizos, Tijmen Blankevoort, Max Welling  
HIGHLIGHT: We show that regularizing the  $\ell_1$ -norm of gradients improves robustness to post-training quantization in neural networks.
- 590, TITLE: SpikeGrad: An ANN-equivalent Computation Model for Implementing Backpropagation with Spikes

- <https://openreview.net/forum?id=rkxs0yHFPH>  
AUTHORS: Johannes C. Thiele, Olivier Bichler, Antoine Dupret  
HIGHLIGHT: An implementation of the backpropagation algorithm using spiking neurons for forward and backward propagation.
- 591, TITLE: On the Relationship between Self-Attention and Convolutional Layers  
<https://openreview.net/forum?id=HJlnC1rKPB>  
AUTHORS: Jean-Baptiste Cordonnier, Andreas Loukas, Martin Jaggi  
HIGHLIGHT: A self-attention layer can perform convolution and often learns to do so in practice.
- 592, TITLE: Learning-Augmented Data Stream Algorithms  
<https://openreview.net/forum?id=HyxJlxBYDH>  
AUTHORS: Tanqiu Jiang, Yi Li, Honghao Lin, Yisong Ruan, David P. Woodruff  
HIGHLIGHT: A self-attention layer can perform convolution and often learns to do so in practice.
- 593, TITLE: Structured Object-Aware Physics Prediction for Video Modeling and Planning  
<https://openreview.net/forum?id=BlE-kxSKDH>  
AUTHORS: Jannik Kossen, Karl Stelzner, Marcel Hussing, Claas Voelcker, Kristian Kersting  
HIGHLIGHT: We propose a structured object-aware video prediction model, which explicitly reasons about objects and demonstrate that it provides high-quality long term video predictions for planning.
- 594, TITLE: Incorporating BERT into Neural Machine Translation  
<https://openreview.net/forum?id=Hyl7ygStwB>  
AUTHORS: Jinhua Zhu, Yingce Xia, Lijun Wu, Di He, Tao Qin, Wengang Zhou, Houqiang Li, Tiejian Liu  
HIGHLIGHT: We propose a structured object-aware video prediction model, which explicitly reasons about objects and demonstrate that it provides high-quality long term video predictions for planning.
- 595, TITLE: MMA Training: Direct Input Space Margin Maximization through Adversarial Training  
<https://openreview.net/forum?id=HkeryxBtPB>  
AUTHORS: Gavin Weiguang Ding, Yash Sharma, Kry Yik Chau Lui, Ruitong Huang  
HIGHLIGHT: We propose MMA training to directly maximize input space margin in order to improve adversarial robustness primarily by removing the requirement of specifying a fixed distortion bound.
- 596, TITLE: Infinite-horizon Off-Policy Policy Evaluation with Multiple Behavior Policies  
<https://openreview.net/forum?id=rkgU1gHtvr>  
AUTHORS: Xinyun Chen, Lu Wang, Yizhe Hang, Heng Ge, Hongyuan Zha  
HIGHLIGHT: A new partially policy-agnostic method for infinite-horizon off-policy policy evaluation with multiple known or unknown behavior policies.
- 597, TITLE: vq-wav2vec: Self-Supervised Learning of Discrete Speech Representations  
<https://openreview.net/forum?id=rylwJxrYDS>  
AUTHORS: Alexei Baevski, Steffen Schneider, Michael Auli  
HIGHLIGHT: Learn how to quantize speech signal and apply algorithms requiring discrete inputs to audio data such as BERT.
- 598, TITLE: Meta-learning curiosity algorithms  
<https://openreview.net/forum?id=BygdyxHFDS>  
AUTHORS: Ferran Alet\*, Martin F. Schneider\*, Tomas Lozano-Perez, Leslie Pack Kaelbling  
HIGHLIGHT: Meta-learning curiosity algorithms by searching through a rich space of programs yields novel mechanisms that generalize across very different reinforcement-learning domains.
- 599, TITLE: Making Efficient Use of Demonstrations to Solve Hard Exploration Problems  
<https://openreview.net/forum?id=SygKycHKDH>  
AUTHORS: Caglar Gulcehre, Tom Le Paine, Bobak Shahriari, Misha Denil, Matt Hoffman, Hubert Soyer, Richard Tanburn, Steven Kapturovski, Neil Rabinowitz, Duncan Williams, Gabriel Barth-Maron, Ziyu Wang, Nando de Freitas, Worlds Team  
HIGHLIGHT: We introduce R2D3, an agent that makes efficient use of demonstrations to solve hard exploration problems in partially observable environments with highly variable initial conditions.
- 600, TITLE: VariBAD: A Very Good Method for Bayes-Adaptive Deep RL via Meta-Learning  
<https://openreview.net/forum?id=Hk19JIBYvr>

- AUTHORS: Luisa Zintgraf, Kyriacos Shiarlis, Maximilian Igl, Sebastian Schulze, Yarin Gal, Katja Hofmann, Shimon Whiteson  
HIGHLIGHT: VariBAD opens a path to tractable approximate Bayes-optimal exploration for deep RL using ideas from meta-learning, Bayesian RL, and approximate variational inference.
- 601, TITLE: Lookahead: A Far-sighted Alternative of Magnitude-based Pruning  
<https://openreview.net/forum?id=ryl3ygHYDB>  
AUTHORS: Sejun Park\*, Jaeho Lee\*, Sangwoo Mo, Jinwoo Shin  
HIGHLIGHT: We study a multi-layer generalization of the magnitude-based pruning.
- 602, TITLE: Spike-based causal inference for weight alignment  
<https://openreview.net/forum?id=rJxWxxSYvB>  
AUTHORS: Jordan Guerguiev, Konrad Kording, Blake Richards  
HIGHLIGHT: We present a learning rule for feedback weights in a spiking neural network that addresses the weight transport problem.
- 603, TITLE: Empirical Bayes Transductive Meta-Learning with Synthetic Gradients  
<https://openreview.net/forum?id=Hkg-xgrYvH>  
AUTHORS: Xu Hu, Pablo Moreno, Yang Xiao, Xi Shen, Guillaume Obozinski, Neil Lawrence  
HIGHLIGHT: We present a learning rule for feedback weights in a spiking neural network that addresses the weight transport problem.
- 604, TITLE: Keep Doing What Worked: Behavior Modelling Priors for Offline Reinforcement Learning  
<https://openreview.net/forum?id=rke7geHtwH>  
AUTHORS: Noah Siegel, Jost Tobias Springenberg, Felix Berkenkamp, Abbas Abdolmaleki, Michael Neunert, Thomas Lampe, Roland Hafner, Nicolas Heess, Martin Riedmiller  
HIGHLIGHT: We develop a method for stable offline reinforcement learning from logged data. The key is to regularize the RL policy towards a learned "advantage weighted" model of the data.
- 605, TITLE: Understanding the Limitations of Conditional Generative Models  
<https://openreview.net/forum?id=r1lPleBFvH>  
AUTHORS: Ethan Fetaya, Joern-Henrik Jacobsen, Will Grathwohl, Richard Zemel  
HIGHLIGHT: We develop a method for stable offline reinforcement learning from logged data. The key is to regularize the RL policy towards a learned "advantage weighted" model of the data.
- 606, TITLE: Demystifying Inter-Class Disentanglement  
<https://openreview.net/forum?id=HyI9xxHYPr>  
AUTHORS: Aviv Gabbay, Yedid Hoshen  
HIGHLIGHT: Latent Optimization for Representation Disentanglement
- 607, TITLE: Mixed-curvature Variational Autoencoders  
<https://openreview.net/forum?id=S1g6xeSKDS>  
AUTHORS: Ondrej Skopek, Gary B?cigneul, Octavian-Eugen Ganea  
HIGHLIGHT: Variational Autoencoders with latent spaces modeled as products of constant curvature Riemannian manifolds improve on image reconstruction over single-manifold variants.
- 608, TITLE: BinaryDuo: Reducing Gradient Mismatch in Binary Activation Network by Coupling Binary Activations  
<https://openreview.net/forum?id=r1x0lxfPS>  
AUTHORS: Hyungjun Kim, Kyungsu Kim, Jinseok Kim, Jae-Joon Kim  
HIGHLIGHT: Variational Autoencoders with latent spaces modeled as products of constant curvature Riemannian manifolds improve on image reconstruction over single-manifold variants.
- 609, TITLE: Model-based reinforcement learning for biological sequence design  
<https://openreview.net/forum?id=HklxgbBKvr>  
AUTHORS: Christof Angermueller, David Dohan, David Belanger, Ramya Deshpande, Kevin Murphy, Lucy Colwell  
HIGHLIGHT: We augment model-free policy learning with a sequence-level surrogate reward functions and count-based visitation bonus and demonstrate effectiveness in the large batch, low-round regime seen in designing DNA and protein sequences.
- 610, TITLE: BayesOpt Adversarial Attack  
<https://openreview.net/forum?id=Hkem-lrtvH>

AUTHORS: Binxin Ru, Adam Cobb, Arno Blaas, Yarin Gal  
HIGHLIGHT: We propose a query-efficient black-box attack which uses Bayesian optimisation in combination with Bayesian model selection to optimise over the adversarial perturbation and the optimal degree of search space dimension reduction.

611, TITLE: Meta Reinforcement Learning with Autonomous Inference of Subtask Dependencies

<https://openreview.net/forum?id=HkgsWxrtPB>

AUTHORS: Sungryull Sohn, Hyunjae Woo, Jongwook Choi, Honglak Lee

HIGHLIGHT: A novel meta-RL method that infers latent subtask structure

612, TITLE: Hypermodels for Exploration

<https://openreview.net/forum?id=ryx6WgStPB>

AUTHORS: Vikranth Dwaracherla, Xiuyuan Lu, Morteza Ibrahimi, Ian Osband, Zheng Wen, Benjamin Van Roy

HIGHLIGHT: Hypermodels can encode posterior distributions similar to large ensembles at much smaller computational cost. This can facilitate significant improvements in exploration.

613, TITLE: RaPP: Novelty Detection with Reconstruction along Projection Pathway

<https://openreview.net/forum?id=HkgeGeBYDB>

AUTHORS: Ki Hyun Kim, Sangwoo Shim, Yongsub Lim, Jongseob Jeon, Jeongwoo Choi, Byungchan Kim, Andre S. Yoon

HIGHLIGHT: A new methodology for novelty detection by utilizing hidden space activation values obtained from a deep autoencoder.

614, TITLE: Dynamics-Aware Embeddings

<https://openreview.net/forum?id=BJgZGeHFPH>

AUTHORS: William Whitney, Rajat Agarwal, Kyunghyun Cho, Abhinav Gupta

HIGHLIGHT: State and action embeddings which incorporate the dynamics improve exploration and RL from pixels.

615, TITLE: Functional Regularisation for Continual Learning with Gaussian Processes

<https://openreview.net/forum?id=HkxCzeHFDB>

AUTHORS: Michalis K. Titsias, Jonathan Schwarz, Alexander G. de G. Matthews, Razvan Pascanu, Yee Whye Teh

HIGHLIGHT: Using inducing point sparse Gaussian process methods to overcome catastrophic forgetting in neural networks.

616, TITLE: You CAN Teach an Old Dog New Tricks! On Training Knowledge Graph Embeddings

<https://openreview.net/forum?id=BkxSmlBFvr>

AUTHORS: Daniel Ruffinelli, Samuel Broscheit, Rainer Gemulla

HIGHLIGHT: We study the impact of training strategies on the performance of knowledge graph embeddings.

617, TITLE: AdvectiveNet: An Eulerian-Lagrangian Fluidic Reservoir for Point Cloud Processing

<https://openreview.net/forum?id=H1eqQeHFDS>

AUTHORS: Xingzhe He, Helen Lu Cao, Bo Zhu

HIGHLIGHT: We present a new grid-particle learning method to process point clouds motivated by computational fluid dynamics.

618, TITLE: Never Give Up: Learning Directed Exploration Strategies

<https://openreview.net/forum?id=Sye57xStvB>

AUTHORS: Adri? Puigdom?nech Badia, Pablo Sprechmann, Alex Vitvitskiy, Daniel Guo, Bilal Piot, Steven Kapturowski, Olivier Tieleman, Martin Arjovsky, Alexander Pritzel, Andrew Bolt, Charles Blundell

HIGHLIGHT: We propose a reinforcement learning agent to solve hard exploration games by learning a range of directed exploratory policies.

619, TITLE: Fair Resource Allocation in Federated Learning

<https://openreview.net/forum?id=ByexElsYDr>

AUTHORS: Tian Li, Maziar Sanjabi, Ahmad Beirami, Virginia Smith

HIGHLIGHT: We propose a novel optimization objective that encourages fairness in heterogeneous federated networks, and develop a scalable method to solve it.

620, TITLE: Smooth markets: A basic mechanism for organizing gradient-based learners

<https://openreview.net/forum?id=B1xMEerYvB>

AUTHORS: David Balduzzi, Wojciech M. Czarnecki, Edward Hughes, Joel Leibo, Ian Gemp, Tom Anthony, Georgios Piliouras, Thore Graepel

HIGHLIGHT: We introduce a class of n-player games suited to gradient-based methods.

- 621, TITLE: StructBERT: Incorporating Language Structures into Pre-training for Deep Language Understanding  
<https://openreview.net/forum?id=BJgQ4ISFPH>  
AUTHORS: Wei Wang, Bin Bi, Ming Yan, Chen Wu, Jiangnan Xia, Zuyi Bao, Liwei Peng, Luo Si  
HIGHLIGHT: We introduce a class of n-player games suited to gradient-based methods.
- 622, TITLE: Training binary neural networks with real-to-binary convolutions  
<https://openreview.net/forum?id=BJg4NgBKvH>  
AUTHORS: Brais Martinez, Jing Yang, Adrian Bulat, Georgios Tzimiropoulos  
HIGHLIGHT: We introduce a class of n-player games suited to gradient-based methods.
- 623, TITLE: Permutation Equivariant Models for Compositional Generalization in Language  
<https://openreview.net/forum?id=SyIVNerFvr>  
AUTHORS: Jonathan Gordon, David Lopez-Paz, Marco Baroni, Diane Bouchacourt  
HIGHLIGHT: We propose a link between permutation equivariance and compositional generalization, and provide equivariant language models  
Based on this hypothesis, we propose a set of tools for constructing equivariant sequence-to-sequence models.
- 624, TITLE: Continual learning with hypernetworks  
<https://openreview.net/forum?id=SJgwNerKvB>  
AUTHORS: Johannes von Oswald, Christian Henning, Jo?o Sacramento, Benjamin F. Grewe  
HIGHLIGHT: We propose a link between permutation equivariance and compositional generalization, and provide equivariant language models
- 625, TITLE: Phase Transitions for the Information Bottleneck in Representation Learning  
<https://openreview.net/forum?id=HJloEIBYvB>  
AUTHORS: Tailin Wu, Ian Fischer  
HIGHLIGHT: We give a theoretical analysis of the Information Bottleneck objective to understand and predict observed phase transitions.
- 626, TITLE: Variational Template Machine for Data-to-Text Generation  
<https://openreview.net/forum?id=HkejNgBtPB>  
AUTHORS: Rong Ye, Wenxian Shi, Hao Zhou, Zhongyu Wei, Lei Li  
HIGHLIGHT: We give a theoretical analysis of the Information Bottleneck objective to understand and predict observed phase transitions.
- 627, TITLE: MEMORY-BASED GRAPH NETWORKS  
<https://openreview.net/forum?id=r1IaNeBYPB>  
AUTHORS: Amir hosein Khasahmadi, Kaveh Hassani, Parsa Moradi, Leo Lee, Quaid Morris  
HIGHLIGHT: We introduce efficient memory layers for graph neural networks
- 628, TITLE: AugMix: A Simple Data Processing Method to Improve Robustness and Uncertainty  
<https://openreview.net/forum?id=S1gmrXHFvB>  
AUTHORS: Dan Hendrycks\*, Norman Mu\*, Ekin Dogus Cubuk, Barret Zoph, Justin Gilmer, Balaji Lakshminarayanan  
HIGHLIGHT: We obtain state-of-the-art on robustness to data shifts, and we maintain calibration under data shift even though even when accuracy drops
- 629, TITLE: AtomNAS: Fine-Grained End-to-End Neural Architecture Search  
<https://openreview.net/forum?id=BylQSxHFwr>  
AUTHORS: Jieru Mei, Yingwei Li, Xiaochen Lian, Xiaojie Jin, Linjie Yang, Alan Yuille, Jianchao Yang  
HIGHLIGHT: A new state-of-the-art on Imagenet for mobile setting
- 630, TITLE: Residual Energy-Based Models for Text Generation  
<https://openreview.net/forum?id=B1I4SgHKDH>  
AUTHORS: Yuntian Deng, Anton Bakhtin, Myle Ott, Arthur Szlam  
HIGHLIGHT: We show that Energy-Based models when trained on the residual of an auto-regressive language model can be used effectively and efficiently to generate text.
- 631, TITLE: A closer look at the approximation capabilities of neural networks

- <https://openreview.net/forum?id=rkevSgrtPr>  
AUTHORS: Kai Fong Ernest Chong  
HIGHLIGHT: A quantitative refinement of the universal approximation theorem via an algebraic approach.
- 632, TITLE: Deep Audio Priors Emerge From Harmonic Convolutional Networks  
<https://openreview.net/forum?id=rygiHxrYDB>  
AUTHORS: Zhoutong Zhang, Yunyun Wang, Chuang Gan, Jiajun Wu, Joshua B. Tenenbaum, Antonio Torralba, William T. Freeman  
HIGHLIGHT: A new operation called Harmonic Convolution makes deep network model audio priors without training.
- 633, TITLE: Expected Information Maximization: Using the I-Projection for Mixture Density Estimation  
<https://openreview.net/forum?id=BygLLHFDS>  
AUTHORS: Philipp Becker, Oleg Arenz, Gerhard Neumann  
HIGHLIGHT: A novel, non-adversarial, approach to learn latent variable models in general and mixture models in particular by computing the I-Projection solely based on samples.
- 634, TITLE: A Meta-Transfer Objective for Learning to Disentangle Causal Mechanisms  
<https://openreview.net/forum?id=ryxWlgbFPS>  
AUTHORS: Yoshua Bengio, Tristan Deleu, Nasim Rahaman, Nan Rosemary Ke, Sebastien Lachapelle, Olexa Bilaniuk, Anirudh Goyal, Christopher Pal  
HIGHLIGHT: This paper proposes a meta-learning objective based on speed of adaptation to transfer distributions to discover a modular decomposition and causal variables.
- 635, TITLE: On the interaction between supervision and self-play in emergent communication  
<https://openreview.net/forum?id=rJxGLlBtWH>  
AUTHORS: Ryan Lowe\*, Abhinav Gupta\*, Jakob Foerster, Douwe Kiela, Joelle Pineau  
HIGHLIGHT: This paper proposes a meta-learning objective based on speed of adaptation to transfer distributions to discover a modular decomposition and causal variables.
- 636, TITLE: Dynamic Model Pruning with Feedback  
<https://openreview.net/forum?id=SJem8ISFwB>  
AUTHORS: Tao Lin, Sebastian U. Stich, Luis Barba, Daniil Dmitriev, Martin Jaggi  
HIGHLIGHT: This paper proposes a meta-learning objective based on speed of adaptation to transfer distributions to discover a modular decomposition and causal variables.
- 637, TITLE: Latent Normalizing Flows for Many-to-Many Cross Domain Mappings  
<https://openreview.net/forum?id=SJxE8erKDH>  
AUTHORS: Shweta Mahajan, Iryna Gurevych, Stefan Roth  
HIGHLIGHT: This paper proposes a meta-learning objective based on speed of adaptation to transfer distributions to discover a modular decomposition and causal variables.
- 638, TITLE: Transferring Optimality Across Data Distributions via Homotopy Methods  
<https://openreview.net/forum?id=S1gElerYwH>  
AUTHORS: Matilde Gargiani, Andrea Zanelli, Quoc Tran Dinh, Moritz Diehl, Frank Hutter  
HIGHLIGHT: We propose a new homotopy-based method to transfer "optimality knowledge" across different data distributions in order to speed up training of deep models.
- 639, TITLE: Regularizing activations in neural networks via distribution matching with the Wasserstein metric  
<https://openreview.net/forum?id=rygwLgrYPB>  
AUTHORS: Taejong Joo, Donggu Kang, Byunghoon Kim  
HIGHLIGHT: We propose a new homotopy-based method to transfer "optimality knowledge" across different data distributions in order to speed up training of deep models.
- 640, TITLE: Mutual Information Gradient Estimation for Representation Learning  
<https://openreview.net/forum?id=ByxaUgrFvH>  
AUTHORS: Liangjian Wen, Yiji Zhou, Lirong He, Mingyuan Zhou, Zenglin Xu  
HIGHLIGHT: We propose a new homotopy-based method to transfer "optimality knowledge" across different data distributions in order to speed up training of deep models.
- 641, TITLE: Efficient Transformer for Mobile Applications

- <https://openreview.net/forum?id=ByeMPIHKPH>  
AUTHORS: Zhanghao Wu\*, Zhijian Liu\*, Ji Lin, Yujun Lin, Song Han  
HIGHLIGHT: We propose a new homotopy-based method to transfer "optimality knowledge" across different data distributions in order to speed up training of deep models.
- 642, TITLE: A Function Space View of Bounded Norm Infinite Width ReLU Nets: The Multivariate Case  
<https://openreview.net/forum?id=H1INPxHKDH>  
AUTHORS: Greg Ongie, Rebecca Willett, Daniel Soudry, Nathan Srebro  
HIGHLIGHT: We characterize the space of functions realizable as a ReLU network with an unbounded number of units (infinite width), but where the Euclidean norm of the weights is bounded.
- 643, TITLE: Adversarial Lipschitz Regularization  
[https://openreview.net/forum?id=Bke\\_DertPB](https://openreview.net/forum?id=Bke_DertPB)  
AUTHORS: D?vid Terj?k  
HIGHLIGHT: alternative to gradient penalty
- 644, TITLE: Compositional Continual Language Learning  
<https://openreview.net/forum?id=rklnDgHtDS>  
AUTHORS: Yuanpeng Li, Liang Zhao, Kenneth Church, Mohamed Elhoseiny  
HIGHLIGHT: alternative to gradient penalty
- 645, TITLE: End to End Trainable Active Contours via Differentiable Rendering  
<https://openreview.net/forum?id=rkxawLHKDr>  
AUTHORS: Shir Gur, Tal Shaharabany, Lior Wolf  
HIGHLIGHT: alternative to gradient penalty
- 646, TITLE: Provable Filter Pruning for Efficient Neural Networks  
<https://openreview.net/forum?id=BJxkOISYDH>  
AUTHORS: Lucas Liebenwein, Cenk Baykal, Harry Lang, Dan Feldman, Daniela Rus  
HIGHLIGHT: A sampling-based filter pruning approach for convolutional neural networks exhibiting provable guarantees on the size and performance of the pruned network.
- 647, TITLE: HOW THE CHOICE OF ACTIVATION AFFECTS TRAINING OF OVERPARAMETRIZED NEURAL NETS  
<https://openreview.net/forum?id=rkgfdeBYvH>  
AUTHORS: Abhishek Panigrahi, Abhishek Shetty, Navin Goyal  
HIGHLIGHT: We provide theoretical results about the effect of activation function on the training of highly overparametrized 2-layer neural networks
- 648, TITLE: Lipschitz constant estimation for Neural Networks via sparse polynomial optimization  
[https://openreview.net/forum?id=rJe4\\_xSFDB](https://openreview.net/forum?id=rJe4_xSFDB)  
AUTHORS: Fabian Latorre, Paul Rolland, Volkan Cevher  
HIGHLIGHT: We provide theoretical results about the effect of activation function on the training of highly overparametrized 2-layer neural networks
- 649, TITLE: State Alignment-based Imitation Learning  
<https://openreview.net/forum?id=rylrdxHFDr>  
AUTHORS: Fangchen Liu, Zhan Ling, Tongzhou Mu, Hao Su  
HIGHLIGHT: We provide theoretical results about the effect of activation function on the training of highly overparametrized 2-layer neural networks
- 650, TITLE: Learning to Group: A Bottom-Up Framework for 3D Part Discovery in Unseen Categories  
<https://openreview.net/forum?id=rk18dIHVvB>  
AUTHORS: Tiange Luo, Kaichun Mo, Zhiao Huang, Siyu Hu, Jiarui Xu, Liwei Wang, Hao Su  
HIGHLIGHT: We propose a learning-based iterative grouping framework which learns a grouping policy to progressively merge small part proposals into bigger ones in a bottom-up fashion and achieve the state-of-the-art performance in open-context setting.
- 651, TITLE: Discriminative Particle Filter Reinforcement Learning for Complex Partial observations  
[https://openreview.net/forum?id=HJ18\\_eHYvS](https://openreview.net/forum?id=HJ18_eHYvS)



AUTHORS: Xiao Ma, Peter Karkus, Nan Ye, David Hsu, Wee Sun Lee  
HIGHLIGHT: We introduce DPFRL, a framework for reinforcement learning under partial and complex observations with a fully differentiable discriminative particle filter

652, TITLE: Unrestricted Adversarial Examples via Semantic Manipulation  
[https://openreview.net/forum?id=Sye\\_OgHFwH](https://openreview.net/forum?id=Sye_OgHFwH)  
AUTHORS: Anand Bhattad, Min Jin Chong, Kaizhao Liang, Bo Li, David Forsyth  
HIGHLIGHT: We introduce unrestricted perturbations that manipulate semantically meaningful image-based visual descriptors -- color and texture -- in order to generate effective and photorealistic adversarial examples.

653, TITLE: Classification-Based Anomaly Detection for General Data  
[https://openreview.net/forum?id=H1IK\\_IBtvS](https://openreview.net/forum?id=H1IK_IBtvS)  
AUTHORS: Liron Bergman, Yedid Hoshen  
HIGHLIGHT: An anomaly detection that: uses random-transformation classification for generalizing to non-image data.

654, TITLE: Scale-Equivariant Steerable Networks  
<https://openreview.net/forum?id=HJgpgrKPS>  
AUTHORS: Ivan Sosnovik, Michal Szmaja, Arnold Smeulders  
HIGHLIGHT: An anomaly detection that: uses random-transformation classification for generalizing to non-image data.

655, TITLE: On Generalization Error Bounds of Noisy Gradient Methods for Non-Convex Learning  
<https://openreview.net/forum?id=SkxxtgHKPS>  
AUTHORS: Jian Li, Xuanyuan Luo, Mingda Qiao  
HIGHLIGHT: We give some generalization error bounds of noisy gradient methods such as SGLD, Langevin dynamics, noisy momentum and so forth.

656, TITLE: Consistency Regularization for Generative Adversarial Networks  
<https://openreview.net/forum?id=S1lxKISKPH>  
AUTHORS: Han Zhang, Zizhao Zhang, Augustus Odena, Honglak Lee  
HIGHLIGHT: We give some generalization error bounds of noisy gradient methods such as SGLD, Langevin dynamics, noisy momentum and so forth.

657, TITLE: Differentiable learning of numerical rules in knowledge graphs  
<https://openreview.net/forum?id=rJleKgrKwS>  
AUTHORS: Po-Wei Wang, Daria Stepanova, Csaba Domokos, J. Zico Kolter  
HIGHLIGHT: We present an efficient approach to integrating numerical comparisons into differentiable rule learning in knowledge graphs

658, TITLE: Learning to Move with Affordance Maps  
<https://openreview.net/forum?id=BJgMFxrYPB>  
AUTHORS: William Qi, Ravi Teja Mullapudi, Saurabh Gupta, Deva Ramanan  
HIGHLIGHT: We address the task of autonomous exploration and navigation using spatial affordance maps that can be learned in a self-supervised manner, these outperform classic geometric baselines while being more sample efficient than contemporary RL algorithms

659, TITLE: Neural tangent kernels, transportation mappings, and universal approximation  
<https://openreview.net/forum?id=HklQYxBKwS>  
AUTHORS: Ziwei Ji, Matus Telgarsky, Ruicheng Xian  
HIGHLIGHT: The NTK linearization is a universal approximator, even when looking arbitrarily close to initialization

660, TITLE: SCALABLE OBJECT-ORIENTED SEQUENTIAL GENERATIVE MODELS  
<https://openreview.net/forum?id=SJxrKgStDH>  
AUTHORS: Jindong Jiang, Sepehr Janghorbani, Gerard De Melo, Sungjin Ahn  
HIGHLIGHT: The NTK linearization is a universal approximator, even when looking arbitrarily close to initialization

661, TITLE: Prediction Poisoning: Towards Defenses Against DNN Model Stealing Attacks  
<https://openreview.net/forum?id=SyevYxHtDB>  
AUTHORS: Tribhuvanesh Orekondy, Bernt Schiele, Mario Fritz  
HIGHLIGHT: We propose the first approach that can resist DNN model stealing/extraction attacks

- 662, TITLE: Domain Adaptive Multiflow Networks  
<https://openreview.net/forum?id=rJxycxHKDS>  
AUTHORS: R?ger Berm?dez-Chac?n, Mathieu Salzmann, Pascal Fua  
HIGHLIGHT: A Multiflow Network is a dynamic architecture for domain adaptation that learns potentially different computational graphs per domain, so as to map them to a common representation where inference can be performed in a domain-agnostic fashion.
- 663, TITLE: Differentiable Programming for Physical Simulation  
<https://openreview.net/forum?id=B1eB5xSFvr>  
AUTHORS: Yuanming Hu, Luke Anderson, Tzu-Mao Li, Qi Sun, Nathan Carr, Jonathan Ragan-Kelley, Fredo Durand  
HIGHLIGHT: We study the problem of learning and optimizing through physical simulations via differentiable programming, using our proposed DiffSim programming language and compiler.
- 664, TITLE: Pitfalls of In-Domain Uncertainty Estimation and Ensembling in Deep Learning  
<https://openreview.net/forum?id=BJxI5gHKDr>  
AUTHORS: Arsenii Ashukha, Alexander Lyzhov, Dmitry Molchanov, Dmitry Vetrov  
HIGHLIGHT: We highlight the problems with common metrics of in-domain uncertainty and perform a broad study of modern ensembling techniques.
- 665, TITLE: Episodic Reinforcement Learning with Associative Memory  
<https://openreview.net/forum?id=HkxjqxBYDB>  
AUTHORS: Guangxiang Zhu\*, Zichuan Lin\*, Guangwen Yang, Chongjie Zhang  
HIGHLIGHT: We highlight the problems with common metrics of in-domain uncertainty and perform a broad study of modern ensembling techniques.
- 666, TITLE: Sub-policy Adaptation for Hierarchical Reinforcement Learning  
<https://openreview.net/forum?id=ByeWogStDS>  
AUTHORS: Alexander Li, Carlos Florensa, Ignasi Clavera, Pieter Abbeel  
HIGHLIGHT: We propose HIPPO, a stable Hierarchical Reinforcement Learning algorithm that can train several levels of the hierarchy simultaneously, giving good performance both in skill discovery and adaptation.
- 667, TITLE: Critical initialisation in continuous approximations of binary neural networks  
<https://openreview.net/forum?id=rylmoxrFDH>  
AUTHORS: George Stamatescu, Federica Gerace, Carlo Lucibello, Ian Fuss, Langford White  
HIGHLIGHT: signal propagation theory applied to continuous surrogates of binary nets; counter intuitive initialisation; reparameterisation trick not helpful
- 668, TITLE: Deep Orientation Uncertainty Learning based on a Bingham Loss  
<https://openreview.net/forum?id=ryloogSKDS>  
AUTHORS: Igor Glitschenski, Wilko Schwarting, Roshni Sahoo, Alexander Amini, Sertac Karaman  
HIGHLIGHT: A method for learning uncertainties over orientations using the Bingham Distribution
- 669, TITLE: Co-Attentive Equivariant Neural Networks: Focusing Equivariance On Transformations Co-Occurring in Data  
<https://openreview.net/forum?id=r1g6ogrtDr>  
AUTHORS: David W. Romero Guzm?n, Mark Hoogendoorn  
HIGHLIGHT: We utilize attention to restrict equivariant neural networks to the set or co-occurring transformations in data.
- 670, TITLE: Mixed Precision DNNs: All you need is a good parametrization  
<https://openreview.net/forum?id=Hyx0slrFvH>  
AUTHORS: Stefan Uhlich, Lukas Mauch, Fabien Pedroni, Kazuki Yoshiyama, Javier Alonso Garcia, Stephen Tiedemann, Thomas Kemp, Akira Nakamura  
HIGHLIGHT: We utilize attention to restrict equivariant neural networks to the set or co-occurring transformations in data.
- 671, TITLE: Information Geometry of Orthogonal Initializations and Training  
<https://openreview.net/forum?id=rkg1ngrFPr>  
AUTHORS: Piotr Aleksander Sok?, Il Memming Park  
HIGHLIGHT: nearly isometric DNN initializations imply low parameter space curvature, and a lower condition number, but that's not always great

- 672, TITLE: Extreme Classification via Adversarial Softmax Approximation  
<https://openreview.net/forum?id=rJxe3xSYDS>  
AUTHORS: Robert Bamler, Stephan Mandt  
HIGHLIGHT: An efficient, unbiased approximation of the softmax loss function for extreme classification
- 673, TITLE: Learning Nearly Decomposable Value Functions Via Communication Minimization  
<https://openreview.net/forum?id=HJx-3grYDB>  
AUTHORS: Tonghan Wang\*, Jianhao Wang\*, Chongyi Zheng, Chongjie Zhang  
HIGHLIGHT: An efficient, unbiased approximation of the softmax loss function for extreme classification
- 674, TITLE: Robust Subspace Recovery Layer for Unsupervised Anomaly Detection  
<https://openreview.net/forum?id=rylb3eBtwr>  
AUTHORS: Chieh-Hsin Lai, Dongmian Zou, Gilad Lerman  
HIGHLIGHT: This work proposes an autoencoder with a novel robust subspace recovery layer for unsupervised anomaly detection and demonstrates state-of-the-art results on various datasets.
- 675, TITLE: Learning to Coordinate Manipulation Skills via Skill Behavior Diversification  
<https://openreview.net/forum?id=ryxB2IBtvH>  
AUTHORS: Youngwoon Lee, Jingyun Yang, Joseph J. Lim  
HIGHLIGHT: We propose to tackle complex tasks of multiple agents by learning composable primitive skills and coordination of the skills.
- 676, TITLE: NAS-BENCH-1SHOT1: BENCHMARKING AND DISSECTING ONE-SHOT NEURAL ARCHITECTURE SEARCH  
<https://openreview.net/forum?id=SJx9ngStPH>  
AUTHORS: Arber Zela, Julien Siems, Frank Hutter  
HIGHLIGHT: We propose to tackle complex tasks of multiple agents by learning composable primitive skills and coordination of the skills.
- 677, TITLE: Conservative Uncertainty Estimation By Fitting Prior Networks  
<https://openreview.net/forum?id=BJlahxHYDS>  
AUTHORS: Kamil Ciosek, Vincent Fortuin, Ryota Tomioka, Katja Hofmann, Richard Turner  
HIGHLIGHT: We provide theoretical support to uncertainty estimates for deep learning obtained fitting random priors.
- 678, TITLE: Understanding Generalization in Recurrent Neural Networks  
<https://openreview.net/forum?id=rkgg6xBYDH>  
AUTHORS: Zhuozhuo Tu, Fengxiang He, Dacheng Tao  
HIGHLIGHT: We provide theoretical support to uncertainty estimates for deep learning obtained fitting random priors.
- 679, TITLE: The Shape of Data: Intrinsic Distance for Data Distributions  
<https://openreview.net/forum?id=HyeblPHYwB>  
AUTHORS: Anton Tsitsulin, Marina Munkhoeva, Davide Mottin, Panagiotis Karras, Alex Bronstein, Ivan Oseledets, Emmanuel Mueller  
HIGHLIGHT: We propose a metric for comparing data distributions based on their geometry while not relying on any positional information.
- 680, TITLE: How to Own the NAS in Your Spare Time  
<https://openreview.net/forum?id=S1erpeBFPB>  
AUTHORS: Sanghyun Hong, Michael Davinroy, Yigitcan Kaya, Dana Dachman-Soled, Tudor Dumitras  
HIGHLIGHT: We design an algorithm that reconstructs the key components of a novel deep learning system by exploiting a small amount of information leakage from a cache side-channel attack, Flush+Reload.
- 681, TITLE: Enabling Deep Spiking Neural Networks with Hybrid Conversion and Spike Timing Dependent Backpropagation  
<https://openreview.net/forum?id=B1xSperKvH>  
AUTHORS: Nitin Rathi, Gopalakrishnan Srinivasan, Priyadarshini Panda, Kaushik Roy  
HIGHLIGHT: We design an algorithm that reconstructs the key components of a novel deep learning system by exploiting a small amount of information leakage from a cache side-channel attack, Flush+Reload.

- 682, TITLE: BREAKING CERTIFIED DEFENSES: SEMANTIC ADVERSARIAL EXAMPLES WITH SPOOFED ROBUSTNESS CERTIFICATES  
<https://openreview.net/forum?id=HJxdTxHYvB>  
AUTHORS: Amin Ghiasi, Ali Shafahi, Tom Goldstein  
HIGHLIGHT: We design an algorithm that reconstructs the key components of a novel deep learning system by exploiting a small amount of information leakage from a cache side-channel attack, Flush+Reload.
- 683, TITLE: Query-efficient Meta Attack to Deep Neural Networks  
<https://openreview.net/forum?id=Skxd6gSYDS>  
AUTHORS: Jiawei Du, Hu Zhang, Joey Tianyi Zhou, Yi Yang, Jiashi Feng  
HIGHLIGHT: We design an algorithm that reconstructs the key components of a novel deep learning system by exploiting a small amount of information leakage from a cache side-channel attack, Flush+Reload.
- 684, TITLE: Massively Multilingual Sparse Word Representations  
<https://openreview.net/forum?id=HyeYTgrFPB>  
AUTHORS: G?bor Berend  
HIGHLIGHT: We propose an efficient algorithm for determining multilingually comparable sparse word representations that we release for 27 typologically diverse languages.
- 685, TITLE: Monotonic Multihead Attention  
<https://openreview.net/forum?id=Hyg96gBKPS>  
AUTHORS: Xutai Ma, Juan Miguel Pino, James Cross, Liezl Puzon, Jiatao Gu  
HIGHLIGHT: Make the transformer streamable with monotonic attention.
- 686, TITLE: Gradients as Features for Deep Representation Learning  
<https://openreview.net/forum?id=BkeoaeHKDS>  
AUTHORS: Fangzhou Mu, Yingyu Liang, Yin Li  
HIGHLIGHT: Given a pre-trained model, we explored the per-sample gradients of the model parameters relative to a task-specific loss, and constructed a linear model that combines gradients of model parameters and the activation of the model.
- 687, TITLE: Pay Attention to Features, Transfer Learn faster CNNs  
<https://openreview.net/forum?id=ryxyCeHtPB>  
AUTHORS: Kafeng Wang, Xitong Gao, Yiren Zhao, Xingjian Li, Dejing Dou, Cheng-Zhong Xu  
HIGHLIGHT: We introduce attentive feature distillation and selection, to fine-tune a large model and produce a faster one.