

- 1, TITLE: Rethinking The Separation Layers In Speech Separation Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414176>  
AUTHORS: Y. Luo, Z. Chen, C. Han, C. Li, T. Zhou and N. Mesgarani  
HIGHLIGHT: In this paper, we empirically examine those questions by designing models with varying configurations in the SIMO and SISO modules.
  
- 2, TITLE: On Permutation Invariant Training For Speech Source Separation  
<https://doi.org/10.1109/ICASSP39728.2021.9413559>  
AUTHORS: X. Liu and J. Pons  
HIGHLIGHT: We study permutation invariant training (PIT), which targets at the permutation ambiguity problem for speaker independent source separation models.
  
- 3, TITLE: Count And Separate: Incorporating Speaker Counting For Continuous Speaker Separation  
<https://doi.org/10.1109/ICASSP39728.2021.9414677>  
AUTHORS: Z. -Q. Wang and D. Wang  
HIGHLIGHT: This study leverages frame-wise speaker counting to switch between speech enhancement and speaker separation for continuous speaker separation.
  
- 4, TITLE: Ultra-Lightweight Speech Separation Via Group Communication  
<https://doi.org/10.1109/ICASSP39728.2021.9414322>  
AUTHORS: Y. Luo, C. Han and N. Mesgarani  
HIGHLIGHT: In this paper, we provide a simple model design paradigm that explicitly designs ultra-lightweight models without sacrificing the performance.
  
- 5, TITLE: Attention Is All You Need In Speech Separation  
<https://doi.org/10.1109/ICASSP39728.2021.9413901>  
AUTHORS: C. Subakan, M. Ravanelli, S. Cornell, M. Bronzi and J. Zhong  
HIGHLIGHT: Transformers are emerging as a natural alternative to standard RNNs, replacing recurrent computations with a multi-head attention mechanism. In this paper, we propose the SepFormer, a novel RNN-free Transformer-based neural network for speech separation.
  
- 6, TITLE: Multichannel Overlapping Speaker Segmentation Using Multiple Hypothesis Tracking Of Acoustic And Spatial Features  
<https://doi.org/10.1109/ICASSP39728.2021.9414130>  
AUTHORS: A. O. T. Hogg, C. Evers and P. A. Naylor  
HIGHLIGHT: In this paper we explore the use of a new multimodal approach for overlapping speaker segmentation that tracks both the fundamental frequency (F0) of the speaker and the speaker's direction of arrival (DOA) simultaneously.
  
- 7, TITLE: Semi-Supervised Singing Voice Separation With Noisy Self-Training  
<https://doi.org/10.1109/ICASSP39728.2021.9413723>  
AUTHORS: Z. Wang, R. Giri, U. Isik, J. -M. Valin and A. Krishnaswamy  
HIGHLIGHT: Given a limited set of labeled data, we present a method to leverage a large volume of unlabeled data to improve the model's performance.
  
- 8, TITLE: Neuro-Steered Music Source Separation With EEG-Based Auditory Attention Decoding And Contrastive-NMF  
<https://doi.org/10.1109/ICASSP39728.2021.9413841>  
AUTHORS: G. Cantisani, S. Essid and G. Richard  
HIGHLIGHT: We propose a novel informed music source separation paradigm, which can be referred to as neuro-steered music source separation.
  
- 9, TITLE: Complex Ratio Masking For Singing Voice Separation  
<https://doi.org/10.1109/ICASSP39728.2021.9414398>  
AUTHORS: Y. Zhang, Y. Liu and D. Wang  
HIGHLIGHT: This paper proposes a complex ratio masking method for voice and accompaniment separation.
  
- 10, TITLE: Transcription Is All You Need: Learning To Separate Musical Mixtures With Score As Supervision  
<https://doi.org/10.1109/ICASSP39728.2021.9413358>  
AUTHORS: Y. -N. Hung, G. Wichern and J. Le Roux  
HIGHLIGHT: In this work, we use musical scores, which are comparatively easy to obtain, as a weak label for training a source separation system.

- 11, TITLE: All For One And One For All: Improving Music Separation By Bridging Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414044>  
AUTHORS: R. Sawata, S. Uhlich, S. Takahashi and Y. Mitsufuji  
HIGHLIGHT: This paper proposes several improvements for music separation with deep neural networks (DNNs), namely a multi-domain loss (MDL) and two combination schemes.
- 12, TITLE: An Hrnet-Blstm Model With Two-Stage Training For Singing Melody Extraction  
<https://doi.org/10.1109/ICASSP39728.2021.9414431>  
AUTHORS: Y. Gao, X. Du, B. Zhu, X. Sun, W. Li and Z. Ma  
HIGHLIGHT: To overcome this problem, we propose to use a pitch refinement method to refine the semitone-level pitch sequences decoded from massive melody MIDI files to generate a large number of fundamental frequency (F0) values for model training.
- 13, TITLE: DeepF0: End-To-End Fundamental Frequency Estimation for Music and Speech Signals  
<https://doi.org/10.1109/ICASSP39728.2021.9414050>  
AUTHORS: S. Singh, R. Wang and Y. Qiu  
HIGHLIGHT: We propose a novel pitch estimation technique called DeepF0, which leverages the available annotated data to directly learns from the raw audio in a data-driven manner.
- 14, TITLE: Differentiable Signal Processing With Black-Box Audio Effects  
<https://doi.org/10.1109/ICASSP39728.2021.9415103>  
AUTHORS: M. A. Mart?nez Ram?rez, O. Wang, P. Smaragdis and N. J. Bryan  
HIGHLIGHT: We present a data-driven approach to automate audio signal processing by incorporating stateful third-party, audio effects as layers within a deep neural network.
- 15, TITLE: Automatic Multitrack Mixing With A Differentiable Mixing Console Of Neural Audio Effects  
<https://doi.org/10.1109/ICASSP39728.2021.9414364>  
AUTHORS: C. J. Steinmetz, J. Pons, S. Pascual and J. Serr?  
HIGHLIGHT: To address these challenges, we propose a domain-inspired model with a strong inductive bias for the mixing task.
- 16, TITLE: Sequence-To-Sequence Singing Voice Synthesis With Perceptual Entropy Loss  
<https://doi.org/10.1109/ICASSP39728.2021.9414348>  
AUTHORS: J. Shi, S. Guo, N. Huo, Y. Zhang and Q. Jin  
HIGHLIGHT: In this work, we propose a Perceptual Entropy (PE) loss derived from a psycho-acoustic hearing model to regularize the network.
- 17, TITLE: Reverb Conversion Of Mixed Vocal Tracks Using An End-To-End Convolutional Deep Neural Network  
<https://doi.org/10.1109/ICASSP39728.2021.9414038>  
AUTHORS: J. Koo, S. Paik and K. Lee  
HIGHLIGHT: In response, we propose an end-to-end system capable of switching the musical reverb factor of two different mixed vocal tracks.
- 18, TITLE: Extending Music Based On Emotion And Tonality Via Generative Adversarial Network  
<https://doi.org/10.1109/ICASSP39728.2021.9413365>  
AUTHORS: B. -W. Tseng, Y. -L. Shen and T. -S. Chi  
HIGHLIGHT: We propose a generative model for music extension in this paper.
- 19, TITLE: Improving The Robustness Of Right Whale Detection In Noisy Conditions Using Denoising Autoencoders And Augmented Training  
<https://doi.org/10.1109/ICASSP39728.2021.9414682>  
AUTHORS: W. Vickers, B. Milner and R. Lee  
HIGHLIGHT: The aim of this paper is to examine denoising autoencoders (DAEs) for improving the detection of right whales recorded in harsh marine environments.
- 20, TITLE: Self-Supervised VQ-VAE for One-Shot Music Style Transfer  
<https://doi.org/10.1109/ICASSP39728.2021.9414235>  
AUTHORS: O. C?fka, A. Ozerov, U. Simsekli and G. Richard  
HIGHLIGHT: In this work, we are specifically interested in the problem of one-shot timbre transfer.

- 21, TITLE: Capturing Temporal Dependencies Through Future Prediction for CNN-Based Audio Classifiers  
<https://doi.org/10.1109/ICASSP39728.2021.9414018>  
AUTHORS: H. Song, J. Han, S. Deng and Z. Du  
HIGHLIGHT: To capture audio temporal dependencies using CNNs, we take a different approach from the purely architecture-induced method and explicitly encode temporal dependencies into the CNN-based audio classifiers.
- 22, TITLE: Segmental Dtw: A Parallelizable Alternative to Dynamic Time Warping  
<https://doi.org/10.1109/ICASSP39728.2021.9413827>  
AUTHORS: T. Tsai  
HIGHLIGHT: In this work we explore parallelizable alternatives to DTW for globally aligning two feature sequences.
- 23, TITLE: Pitch-Timbre Disentanglement Of Musical Instrument Sounds Based On Vae-Based Metric Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9414059>  
AUTHORS: K. Tanaka, R. Nishikimi, Y. Bando, K. Yoshii and S. Morishima  
HIGHLIGHT: This paper describes a representation learning method for disentangling an arbitrary musical instrument sound into latent pitch and timbre representations.
- 24, TITLE: Asynchronous Acoustic Echo Cancellation Over Wireless Channels  
<https://doi.org/10.1109/ICASSP39728.2021.9414288>  
AUTHORS: R. Ayrapietian, P. Hilmes, M. Mansour, T. Kristjansson and C. Murgia  
HIGHLIGHT: We introduce a novel acoustic echo cancellation framework for systems where the loudspeaker and the microphone array are not synchronized.
- 25, TITLE: Combining Adaptive Filtering And Complex-Valued Deep Postfiltering For Acoustic Echo Cancellation  
<https://doi.org/10.1109/ICASSP39728.2021.9414868>  
AUTHORS: M. M. Halimeh, T. Haubner, A. Briegleb, A. Schmidt and W. Kellermann  
HIGHLIGHT: In this contribution, we introduce a novel approach to noise-robust acoustic echo cancellation employing a complex-valued Deep Neural Network (DNN) for postfiltering.
- 26, TITLE: Deep Residual Echo Suppression With A Tunable Tradeoff Between Signal Distortion And Echo Suppression  
<https://doi.org/10.1109/ICASSP39728.2021.9414958>  
AUTHORS: A. Ivry, I. Cohen and B. Berdugo  
HIGHLIGHT: In this paper, we propose a residual echo suppression method using a UNet neural network that directly maps the outputs of a linear acoustic echo canceler to the desired signal in the spectral domain.
- 27, TITLE: Robust STFT Domain Multi-Channel Acoustic Echo Cancellation with Adaptive Decorrelation of the Reference Signals  
<https://doi.org/10.1109/ICASSP39728.2021.9413836>  
AUTHORS: S. Bagheri and D. Giacobello  
HIGHLIGHT: In this paper, we present an algorithm for multi-channel acoustic echo cancellation for a high-fidelity audio reproduction system equipped with a microphone array for voice control.
- 28, TITLE: A Method for Determining Periodically Time-Varying Bias and Its Applications in Acoustic Feedback Cancellation  
<https://doi.org/10.1109/ICASSP39728.2021.9414685>  
AUTHORS: M. Guo  
HIGHLIGHT: In this work, we make use of that knowledge and propose a method to detect different acoustic situations, based on the level of residual bias.
- 29, TITLE: Weighted Recursive Least Square Filter and Neural Network Based Residual ECHO Suppression for the AEC-Challenge  
<https://doi.org/10.1109/ICASSP39728.2021.9414623>  
AUTHORS: Z. Wang, Y. Na, Z. Liu, B. Tian and Q. Fu  
HIGHLIGHT: This paper presents a real-time Acoustic Echo Cancellation (AEC) algorithm submitted to the AEC-Challenge.
- 30, TITLE: ICASSP 2021 Acoustic Echo Cancellation Challenge: Integrated Adaptive Echo Cancellation with Time Alignment and Deep Learning-Based Residual Echo Plus Noise Suppression  
<https://doi.org/10.1109/ICASSP39728.2021.9414462>  
AUTHORS: R. Peng, L. Cheng, C. Zheng and X. Li  
HIGHLIGHT: This paper describes a three-stage acoustic echo cancellation (AEC) and suppression framework for the ICASSP 2021 AEC Challenge.

- 31, TITLE: ICASSP 2021 Acoustic Echo Cancellation Challenge: Datasets, Testing Framework, and Results  
<https://doi.org/10.1109/ICASSP39728.2021.9413457>  
AUTHORS: K. Sridhar et al.  
HIGHLIGHT: In this challenge, we open source two large datasets to train AEC models under both single talk and double talk scenarios.
- 32, TITLE: AEC in A Netshell: on Target and Topology Choices for FCRN Acoustic Echo Cancellation  
<https://doi.org/10.1109/ICASSP39728.2021.9414715>  
AUTHORS: J. Franzen, E. Seidel and T. Fingscheidt  
HIGHLIGHT: In this work we will heal this issue and significantly improve the near-end speech component quality over existing approaches.
- 33, TITLE: Kernel-Interpolation-Based Filtered-X Least Mean Square for Spatial Active Noise Control In Time Domain  
<https://doi.org/10.1109/ICASSP39728.2021.9414842>  
AUTHORS: J. Brunnström and S. Koyama  
HIGHLIGHT: Time-domain spatial active noise control (ANC) algorithms based on kernel interpolation of a sound field are proposed.
- 34, TITLE: Wave-Domain Optimization of Secondary Source Placement Free From Information of Error Sensor Positions  
<https://doi.org/10.1109/ICASSP39728.2021.9414622>  
AUTHORS: J. Xu, K. Chen and Y. Li  
HIGHLIGHT: In this study, a method free from the information of specific error sensors positions is proposed.
- 35, TITLE: Lasaft: Latent Source Attentive Frequency Transformation For Conditioned Source Separation  
<https://doi.org/10.1109/ICASSP39728.2021.9413896>  
AUTHORS: W. Choi, M. Kim, J. Chung and S. Jung  
HIGHLIGHT: The goal of this paper is to extend the FT block to fit the multi-source task.
- 36, TITLE: Surrogate Source Model Learning for Determined Source Separation  
<https://doi.org/10.1109/ICASSP39728.2021.9414255>  
AUTHORS: R. Scheibler and M. Togami  
HIGHLIGHT: We propose to learn surrogate functions of universal speech priors for determined blind speech separation.
- 37, TITLE: Auditory Filterbanks Benefit Universal Sound Source Separation  
<https://doi.org/10.1109/ICASSP39728.2021.9414105>  
AUTHORS: H. Li, K. Chen and B. U. Seeber  
HIGHLIGHT: We proposed parameterized Gammatone and Gammachirp filterbanks, which improved performance with fewer parameters and better interpretability.
- 38, TITLE: What's all the Fuss about Free Universal Sound Separation Data?  
<https://doi.org/10.1109/ICASSP39728.2021.9414774>  
AUTHORS: S. Wisdom et al.  
HIGHLIGHT: We introduce the Free Universal Sound Separation (FUSS) dataset, a new corpus for experiments in separating mixtures of an unknown number of sounds from an open domain of sound types.
- 39, TITLE: SepNet: A Deep Separation Matrix Prediction Network for Multichannel Audio Source Separation  
<https://doi.org/10.1109/ICASSP39728.2021.9414884>  
AUTHORS: S. Inoue, H. Kameoka, L. Li and S. Makino  
HIGHLIGHT: In this paper, we propose SepNet, a deep neural network (DNN) designed to predict separation matrices from multichannel observations.
- 40, TITLE: CDPAM: Contrastive Learning for Perceptual Audio Similarity  
<https://doi.org/10.1109/ICASSP39728.2021.9413711>  
AUTHORS: P. Manocha, Z. Jin, R. Zhang and A. Finkelstein  
HIGHLIGHT: This paper introduces CDPAM ?a metric that builds on and advances DPAM.
- 41, TITLE: Linear Multichannel Blind Source Separation based on Time-Frequency Mask Obtained by Harmonic/Percussive Sound Separation

<https://doi.org/10.1109/ICASSP39728.2021.9413494>

AUTHORS: S. Oyabu, D. Kitamura and K. Yatabe

HIGHLIGHT: Building up on this framework, in this paper, we propose a unification of determined BSS and harmonic/percussive sound separation (HPSS).

42, TITLE: Multichannel-based Learning for Audio Object Extraction

<https://doi.org/10.1109/ICASSP39728.2021.9414585>

AUTHORS: D. Arteaga and J. Pons

HIGHLIGHT: Here, we propose a novel deep learning approach to object extraction that learns from the multichannel renders of object-based productions, instead of directly learning from the audio objects themselves.

43, TITLE: DBnet: Doa-Driven Beamforming Network for end-to-end Reverberant Sound Source Separation

<https://doi.org/10.1109/ICASSP39728.2021.9414187>

AUTHORS: A. Aroudi and S. Braun

HIGHLIGHT: In this paper we propose a direction-of-arrival-driven beamforming network (DBnet) consisting of direction-of-arrival (DOA) estimation and beamforming layers for end-to-end source separation.

44, TITLE: Joint Dereverberation and Separation With Iterative Source Steering

<https://doi.org/10.1109/ICASSP39728.2021.9413478>

AUTHORS: T. Nakashima, R. Scheibler, M. Togami and N. Ono

HIGHLIGHT: We propose a new algorithm for joint dereverberation and blind source separation (DR-BSS).

45, TITLE: Exploiting Non-Negative Matrix Factorization for Binaural Sound Localization in the Presence of Directional Interference

<https://doi.org/10.1109/ICASSP39728.2021.9414233>

AUTHORS: I. ?rnolfsson, T. Dau, N. Ma and T. May

HIGHLIGHT: This study presents a novel solution to the problem of binaural localization of a speaker in the presence of interfering directional noise and reverberation.

46, TITLE: Blind Extraction of Moving Audio Source in a Challenging Environment Supported by Speaker Identification Via X-Vectors

<https://doi.org/10.1109/ICASSP39728.2021.9414331>

AUTHORS: J. Malek, J. Jansky, T. Kounovsky, Z. Koldovsky and J. Zdansky

HIGHLIGHT: We propose a novel approach for semi-supervised extraction of a moving audio source of interest (SOI) applicable in reverberant and noisy environments.

47, TITLE: Mind the Beat: Detecting Audio Onsets from EEG Recordings of Music Listening

<https://doi.org/10.1109/ICASSP39728.2021.9414245>

AUTHORS: A. Vinay, A. Lerch and G. Leslie

HIGHLIGHT: We propose a deep learning approach to predicting audio event onsets in electroencephalogram (EEG) recorded from users as they listen to music.

48, TITLE: Don't Look Back: An Online Beat Tracking Method Using RNN and Enhanced Particle Filtering

<https://doi.org/10.1109/ICASSP39728.2021.9413915>

AUTHORS: M. Heydari and Z. Duan

HIGHLIGHT: We propose Don't Look back! (DLB), a novel approach optimized for efficiency when performing OBT.

49, TITLE: Singing Melody Extraction from Polyphonic Music based on Spectral Correlation Modeling

<https://doi.org/10.1109/ICASSP39728.2021.9414190>

AUTHORS: X. Du, B. Zhu, Q. Kong and Z. Ma

HIGHLIGHT: In this paper, we explore the idea of modeling spectral correlation explicitly for melody extraction.

50, TITLE: Improving Automatic Drum Transcription Using Large-Scale Audio-to-Midi Aligned Data

<https://doi.org/10.1109/ICASSP39728.2021.9414409>

AUTHORS: I. -C. Wei, C. -W. Wu and L. Su

HIGHLIGHT: To tackle this issue, we propose a semi-automatic way of compiling a labeled dataset using the audio-to-MIDI alignment technique.

51, TITLE: Frequency-Temporal Attention Network for Singing Melody Extraction

<https://doi.org/10.1109/ICASSP39728.2021.9413444>

AUTHORS: S. Yu, X. Sun, Y. Yu and W. Li  
HIGHLIGHT: Inspired by these intrinsic characteristics, a frequency-temporal attention network is proposed to mimic human auditory for singing melody extraction.

52, TITLE: Statistical Correction of Transcribed Melody Notes Based on Probabilistic Integration of a Music Language Model and a Transcription Error Model  
<https://doi.org/10.1109/ICASSP39728.2021.9414249>  
AUTHORS: Y. Hiramatsu, G. Shibata, R. Nishikimi, E. Nakamura and K. Yoshii  
HIGHLIGHT: This paper describes a statistical post-processing method for automatic singing transcription that corrects pitch and rhythm errors included in a transcribed note sequence.

53, TITLE: Reliability Assessment of Singing Voice F0-Estimates Using Multiple Algorithms  
<https://doi.org/10.1109/ICASSP39728.2021.9413372>  
AUTHORS: S. Rosenzweig, F. Scherbaum and M. M?ller  
HIGHLIGHT: In this work, we consider an approach to automatically assess the reliability of F0-trajectories estimated from monophonic singing voice recordings.

54, TITLE: End-to-End Lyrics Recognition with Voice to Singing Style Transfer  
<https://doi.org/10.1109/ICASSP39728.2021.9415096>  
AUTHORS: S. Basak, S. Agarwal, S. Ganapathy and N. Takahashi  
HIGHLIGHT: In this paper, we propose a data augmentation method that converts natural speech to singing voice based on vocoder based speech synthesizer.

55, TITLE: Singing Language Identification Using a Deep Phonotactic Approach  
<https://doi.org/10.1109/ICASSP39728.2021.9414203>  
AUTHORS: L. Renault, A. Vaglio and R. Hennequin  
HIGHLIGHT: This work presents a modernized phonotactic system for SLID on polyphonic music: phoneme recognition is performed with a Connectionist Temporal Classification (CTC)-based acoustic model trained with multilingual data, before language classification with a recurrent model based on the phonemes estimation.

56, TITLE: On the Preparation and Validation of a Large-Scale Dataset of Singing Transcription  
<https://doi.org/10.1109/ICASSP39728.2021.9414601>  
AUTHORS: J. -Y. Wang and J. -S. R. Jang  
HIGHLIGHT: This paper proposes a large-scale dataset for singing transcription, along with some methods for fine-tuning and validating its contents.

57, TITLE: Joint Multi-Pitch Detection and Score Transcription for Polyphonic Piano Music  
<https://doi.org/10.1109/ICASSP39728.2021.9413601>  
AUTHORS: L. Liu, V. Morfi and E. Benetos  
HIGHLIGHT: In this paper, we propose a method for joint multi-pitch detection and score transcription for polyphonic piano music.

58, TITLE: Karaoke Key Recommendation Via Personalized Competence-Based Rating Prediction  
<https://doi.org/10.1109/ICASSP39728.2021.9414524>  
AUTHORS: Y. Wang, S. Tanaka, K. Yokoyama, H. -T. Wu and Y. Fang  
HIGHLIGHT: In this paper, we address a novel task of recommending a suitable key for a user to sing a given song to meet his or her vocal competence, by proposing the Personalized Competence-based Rating Prediction (PCRP) model.

59, TITLE: A Closed-Loop Gain-Control Feedback Model for The Medial Efferent System of The Descending Auditory Pathway  
<https://doi.org/10.1109/ICASSP39728.2021.9414195>  
AUTHORS: A. Farhadi, S. G. Jennings, E. A. Strickland and L. H. Carney  
HIGHLIGHT: We have implemented a dynamic, closed-loop gain-control system into an existing auditory model to simulate parts of the efferent system.

60, TITLE: DHASP: Differentiable Hearing Aid Speech Processing  
<https://doi.org/10.1109/ICASSP39728.2021.9414571>  
AUTHORS: Z. Tu, N. Ma and J. Barker  
HIGHLIGHT: In this paper, we explore an alternative approach to finding the optimal fitting by introducing a hearing aid speech processing framework, in which the fitting is optimised in an automated way using an intelligibility objective function based on the HASPI physiological auditory model.

- 61, TITLE: Computationally Efficient DNN-Based Approximation of an Auditory Model for Applications in Speech Processing  
https://doi.org/10.1109/ICASSP39728.2021.9413993  
AUTHORS: A. Nagathil, F. G?bel, A. Nelus and I. C. Bruce  
HIGHLIGHT: Hence, in this work we propose and evaluate DNN-based approximations of a state-of-the-art auditory model.
- 62, TITLE: Cascaded All-Pass Filters with Randomized Center Frequencies and Phase Polarity for Acoustic and Speech Measurement and Data Augmentation  
https://doi.org/10.1109/ICASSP39728.2021.9415057  
AUTHORS: H. Kawahara and K. Yatabe  
HIGHLIGHT: We introduce a new member of TSP (Time Stretched Pulse) for acoustic and speech measurement infrastructure, based on a simple all-pass filter and systematic randomization.
- 63, TITLE: Probing Acoustic Representations for Phonetic Properties  
https://doi.org/10.1109/ICASSP39728.2021.9414776  
AUTHORS: D. Ma, N. Ryant and M. Liberman  
HIGHLIGHT: We compare features from two conventional and four pre-trained systems in some simple frame-level phonetic classification tasks, with classifiers trained on features from one version of the TIMIT dataset and tested on features from another.
- 64, TITLE: An End-To-End Non-Intrusive Model for Subjective and Objective Real-World Speech Assessment Using a Multi-Task Framework  
https://doi.org/10.1109/ICASSP39728.2021.9414182  
AUTHORS: Z. Zhang, P. Vyas, X. Dong and D. S. Williamson  
HIGHLIGHT: In this paper, we propose a novel multi-task non-intrusive approach that is capable of simultaneously estimating both subjective and objective scores of real-world speech, to help facilitate learning.
- 65, TITLE: Few-Shot Continual Learning for Audio Classification  
https://doi.org/10.1109/ICASSP39728.2021.9413584  
AUTHORS: Y. Wang, N. J. Bryan, M. Cartwright, J. Pablo Bello and J. Salamon  
HIGHLIGHT: In this work, we introduce a few-shot continual learning framework for audio classification, where we can continuously expand a trained base classifier to recognize novel classes based on only few labeled data at inference time.
- 66, TITLE: Zero-Shot Audio Classification with Factored Linear and Nonlinear Acoustic-Semantic Projections  
https://doi.org/10.1109/ICASSP39728.2021.9414994  
AUTHORS: H. Xie, O. R?s?nen and T. Virtanen  
HIGHLIGHT: In this paper, we study zero-shot learning in audio classification through factored linear and nonlinear acoustic-semantic projections between audio instances and sound classes.
- 67, TITLE: Unsupervised and Semi-Supervised Few-Shot Acoustic Event Classification  
https://doi.org/10.1109/ICASSP39728.2021.9414546  
AUTHORS: H. -P. Huang, K. C. Puvvada, M. Sun and C. Wang  
HIGHLIGHT: Here, we study unsupervised and semi-supervised learning approaches for few-shot AEC.
- 68, TITLE: Flow-Based Self-Supervised Density Estimation for Anomalous Sound Detection  
https://doi.org/10.1109/ICASSP39728.2021.9414662  
AUTHORS: K. Dohi, T. Endo, H. Purohit, R. Tanabe and Y. Kawaguchi  
HIGHLIGHT: To develop a machine sound monitoring system, a method for detecting anomalous sound is proposed.
- 69, TITLE: Self-Training for Sound Event Detection in Audio Mixtures  
https://doi.org/10.1109/ICASSP39728.2021.9414450  
AUTHORS: S. Park, A. Bellur, D. K. Han and M. Elhilali  
HIGHLIGHT: In order to address limitations in availability of training data, this work proposes a self-training technique to leverage unlabeled datasets in supervised learning using pseudo label estimation.
- 70, TITLE: Prototypical Networks for Domain Adaptation in Acoustic Scene Classification  
https://doi.org/10.1109/ICASSP39728.2021.9414876  
AUTHORS: S. Singh, H. L. Bear and E. Benetos



**HIGHLIGHT:** In the search for an optimal solution to the said problem, we explore a metric learning approach called prototypical networks using the TUT Urban Acoustic Scenes dataset, which consists of 10 different acoustic scenes recorded across 10 cities.

71, **TITLE:** A Global-Local Attention Framework for Weakly Labelled Audio Tagging

<https://doi.org/10.1109/ICASSP39728.2021.9414357>

**AUTHORS:** H. Wang, Y. Zou and W. Wang

**HIGHLIGHT:** To address this issue, we propose a novel two-stream framework for audio tagging by exploiting the global and local information of sound events.

72, **TITLE:** An Improved Mean Teacher Based Method for Large Scale Weakly Labeled Semi-Supervised Sound Event Detection

<https://doi.org/10.1109/ICASSP39728.2021.9414931>

**AUTHORS:** X. Zheng, Y. Song, I. McLoughlin, L. Liu and L. -R. Dai

**HIGHLIGHT:** This paper presents an improved mean teacher (MT) based method for large-scale weakly labeled semi-supervised sound event detection (SED), by focusing on learning a better student model.

73, **TITLE:** Comparison of Deep Co-Training and Mean-Teacher Approaches for Semi-Supervised Audio Tagging

<https://doi.org/10.1109/ICASSP39728.2021.9415116>

**AUTHORS:** L. Cances and T. Pellegrini

**HIGHLIGHT:** In this work, we adapted the Deep-Co-Training algorithm (DCT) to perform AT, and compared it to another SSL approach called Mean Teacher (MT), that has been used by the winning participants of the DCASE competitions these last two years.

74, **TITLE:** The Benefit of Temporally-Strong Labels in Audio Event Classification

<https://doi.org/10.1109/ICASSP39728.2021.9414579>

**AUTHORS:** S. Hershey et al.

**HIGHLIGHT:** To reveal the importance of temporal precision in ground truth audio event labels, we collected precise (~0.1 sec resolution) "strong" labels for a portion of the AudioSet dataset.

75, **TITLE:** Unsupervised Contrastive Learning of Sound Event Representations

<https://doi.org/10.1109/ICASSP39728.2021.9415009>

**AUTHORS:** E. Fonseca, D. Ortego, K. McGuinness, N. E. O'Connor and X. Serra

**HIGHLIGHT:** In this work, we explore unsupervised contrastive learning as a way to learn sound event representations.

76, **TITLE:** Sound Event Detection by Consistency Training and Pseudo-Labeling With Feature-Pyramid Convolutional Recurrent Neural Networks

<https://doi.org/10.1109/ICASSP39728.2021.9414350>

**AUTHORS:** C. -Y. Koh, Y. -S. Chen, Y. -W. Liu and M. R. Bai

**HIGHLIGHT:** To exploit large amount of unlabeled in-domain data efficiently, we applied three semi-supervised learning strategies: interpolation consistency training (ICT), shift consistency training (SCT), and weakly pseudo-labeling.

77, **TITLE:** SESQA: Semi-Supervised Learning for Speech Quality Assessment

<https://doi.org/10.1109/ICASSP39728.2021.9414052>

**AUTHORS:** J. Serr?, J. Pons and S. Pascual

**HIGHLIGHT:** In this work, we tackle these problems with a semi-supervised learning approach, combining available annotations with programmatically generated data, and using 3 different optimization criteria together with 5 complementary auxiliary tasks.

78, **TITLE:** Detecting Signal Corruptions in Voice Recordings For Speech Therapy

<https://doi.org/10.1109/ICASSP39728.2021.9414383>

**AUTHORS:** H. Nyl?n, S. Chatterjee and S. Ternstr?m

**HIGHLIGHT:** In this article we design an experimental setup to detect disturbances in voice recordings, such as additive noise, clipping, infrasound and random muting.

79, **TITLE:** MBNET: MOS Prediction for Synthesized Speech with Mean-Bias Network

<https://doi.org/10.1109/ICASSP39728.2021.9413877>

**AUTHORS:** Y. Leng, X. Tan, S. Zhao, F. Soong, X. -Y. Li and T. Qin

**HIGHLIGHT:** In this paper, we propose MBNet, a MOS predictor with a mean subnet and a bias subnet to better utilize every judge score in MOS datasets, where the mean subnet is used to predict the mean score of each utterance similar to that in previous



works, and the bias subnet to predict the bias score (the difference between the mean score and each individual judge score) and capture the personal preference of individual judges.

80, TITLE: Non-Intrusive Binaural Prediction of Speech Intelligibility Based on Phoneme Classification

<https://doi.org/10.1109/ICASSP39728.2021.9413874>

AUTHORS: J. Roßbach, S. Rüttges, C. F. Hauth, T. Brand and B. T. Meyer

HIGHLIGHT: In this study, we explore an approach for modeling speech intelligibility in spatial acoustic scenes.

81, TITLE: Warp-Q: Quality Prediction for Generative Neural Speech Codecs

<https://doi.org/10.1109/ICASSP39728.2021.9414901>

AUTHORS: W. A. Jassim, J. Skoglund, M. Chinen and A. Hines

HIGHLIGHT: We present WARP-Q, a full-reference objective speech quality metric that uses dynamic time warping cost for MFCC speech representations.

82, TITLE: Crowdsourcing Approach for Subjective Evaluation of Echo Impairment

<https://doi.org/10.1109/ICASSP39728.2021.9414904>

AUTHORS: R. Cutler, B. Nadari, M. Loide, S. Sootla and A. Saabas

HIGHLIGHT: We then introduce an open-source crowdsourcing approach for subjective evaluation of echo impairment which can be used to evaluate the performance of AECs.

83, TITLE: Amplitude Matching: Majorization/Minimization Algorithm for Sound Field Control Only with Amplitude Constraint

<https://doi.org/10.1109/ICASSP39728.2021.9414855>

AUTHORS: S. Koyama, T. Amakasu, N. Ueno and H. Saruwatari

HIGHLIGHT: A sound field control method for synthesizing a desired amplitude distribution inside a target region, amplitude matching, is proposed.

84, TITLE: 3D Multizone Soundfield Reproduction in a Reverberant Environment Using Intensity Matching Method

<https://doi.org/10.1109/ICASSP39728.2021.9414077>

AUTHORS: H. Zuo, T. D. Abhayapala and P. N. Samarasinghe

HIGHLIGHT: We address this challenge and propose a multizone reproduction method for 3D soundfield in a reverberant room based on intensity matching.

85, TITLE: The Far-Field Equatorial Array for Binaural Rendering

<https://doi.org/10.1109/ICASSP39728.2021.9414368>

AUTHORS: J. Ahrens, H. Helmholtz, D. L. Alon and S. V. A. Gar

HIGHLIGHT: We present a method for obtaining a spherical harmonic representation of a sound field based on a microphone array along the equator of a rigid spherical scatterer.

86, TITLE: Spherical Harmonic Representation for Dynamic Sound-Field Measurements

<https://doi.org/10.1109/ICASSP39728.2021.9413708>

AUTHORS: F. Katzberg, M. Maass and A. Mertins

HIGHLIGHT: In this paper, we present a new physical interpretation of the dynamic sampling problem.

87, TITLE: Direction Preserving Wind Noise Reduction Of B-Format Signals

<https://doi.org/10.1109/ICASSP39728.2021.9415089>

AUTHORS: A. Herzog, D. Mirabilii and E. A. P. Habets

HIGHLIGHT: In this work, methods to reduce wind noise while limiting the spatial distortions of the original signal are proposed based on recent works of the present authors.

88, TITLE: Refinement of Direction of Arrival Estimators by Majorization-Minimization Optimization on the Array Manifold

<https://doi.org/10.1109/ICASSP39728.2021.9414798>

AUTHORS: R. Scheibler and M. Togami

HIGHLIGHT: Unlike most conventional methods that rely exclusively on grid search, we introduce a continuous optimization algorithm to refine DOA estimates beyond the resolution of the initial grid.

89, TITLE: On the Predictability of Hrtfs from Ear Shapes Using Deep Networks

<https://doi.org/10.1109/ICASSP39728.2021.9414042>

AUTHORS: Y. Zhou, H. Jiang and V. K. Ithapu

**HIGHLIGHT:** Using 3D ear shapes as inputs, we explore the bounds of HRTF predictability using deep neural networks.

90, **TITLE:** Applied Methods for Sparse Sampling of Head-Related Transfer Functions

<https://doi.org/10.1109/ICASSP39728.2021.9413976>

**AUTHORS:** L. Arbel, Z. Ben-Hur, D. L. Alon and B. Rafaely

**HIGHLIGHT:** This paper describes the application of two methods for ear-aligned HRTF interpolation by sparse sampling: Orthogonal Matching Pursuit and Principal Component Analysis.

91, **TITLE:** Personalized HRTF Modeling Using DNN-Augmented BEM

<https://doi.org/10.1109/ICASSP39728.2021.9414448>

**AUTHORS:** M. Zhang, J. -H. Wang and D. L. James

**HIGHLIGHT:** In this paper, we propose a new deep learning method that combines measurements and numerical simulations to take the best of three worlds.

92, **TITLE:** Efficient Training Data Generation for Phase-Based DOA Estimation

<https://doi.org/10.1109/ICASSP39728.2021.9414070>

**AUTHORS:** F. H?bner, W. Mack and E. A. P. Habets

**HIGHLIGHT:** We propose a low complexity online data generation method to train DL models with a phase-based feature input.

93, **TITLE:** Acoustic Reflectors Localization from Stereo Recordings Using Neural Networks

<https://doi.org/10.1109/ICASSP39728.2021.9414473>

**AUTHORS:** G. Bologni, R. Heusdens and J. Martinez

**HIGHLIGHT:** We propose a fully convolutional network (FCN) that localizes reflective surfaces under the relaxed assumptions that (i) a compact array of only two microphones is available, (ii) emitter and receivers are not synchronized, and (iii) both the excitation signals and the impulse responses of the enclosures are unknown.

94, **TITLE:** Detecting Acoustic Reflectors Using A Robot's Ego-Noise

<https://doi.org/10.1109/ICASSP39728.2021.9414061>

**AUTHORS:** U. Saqib, A. Deleforge and J. R. Jensen

**HIGHLIGHT:** In this paper, we propose a method to estimate the proximity of an acoustic reflector, e.g., a wall, using ego-noise, i.e., the noise produced by the moving parts of a listening robot.

95, **TITLE:** Prediction of Object Geometry from Acoustic Scattering Using Convolutional Neural Networks

<https://doi.org/10.1109/ICASSP39728.2021.9414743>

**AUTHORS:** Z. Fan, V. Vineet, C. Lu, T. W. Wu and K. McMullen

**HIGHLIGHT:** The present work proposes a method to infer object geometry from scattering features by training convolutional neural networks.

96, **TITLE:** Blind Amplitude Estimation of Early Room Reflections Using Alternating Least Squares

<https://doi.org/10.1109/ICASSP39728.2021.9413873>

**AUTHORS:** T. Shlomo and B. Rafaely

**HIGHLIGHT:** This work presents a preliminary attempt to blindly estimate reflection amplitudes.

97, **TITLE:** Acoustic Analysis and Dataset of Transitions Between Coupled Rooms

<https://doi.org/10.1109/ICASSP39728.2021.9415122>

**AUTHORS:** T. McKenzie, S. J. Schlecht and V. Pulkki

**HIGHLIGHT:** This paper presents the measurement and analysis of a dataset of spatial room impulse responses for the transition between four coupled room pairs.

98, **TITLE:** On Loss Functions for Deep-Learning Based T60 Estimation

<https://doi.org/10.1109/ICASSP39728.2021.9414442>

**AUTHORS:** Y. Li, Y. Liu and D. S. Williamson

**HIGHLIGHT:** In this paper, we propose a composite classification- and regression-based cost function for training a deep neural network that predicts T60 for a variety of reverberant signals.

99, **TITLE:** Towards Listening to 10 People Simultaneously: An Efficient Permutation Invariant Training of Audio Source Separation Using Sinkhorn's Algorithm

<https://doi.org/10.1109/ICASSP39728.2021.9414508>

**AUTHORS:** H. Tachibana

**HIGHLIGHT:** To overcome this limitation, this paper proposes a SinkPIT, a novel variant of the PIT losses, which is much more efficient than the ordinary PIT loss when N is large.

100, **TITLE:** Accelerating Auxiliary Function-Based Independent Vector Analysis

<https://doi.org/10.1109/ICASSP39728.2021.9413950>

**AUTHORS:** A. Brendel and W. Kellermann

**HIGHLIGHT:** To this end, we investigate techniques which accelerate the convergence of the AuxIVA update rules without extra computational cost.

101, **TITLE:** One-Shot Conditional Audio Filtering of Arbitrary Sounds

<https://doi.org/10.1109/ICASSP39728.2021.9414003>

**AUTHORS:** B. Gfeller, D. Roblek and M. Tagliasacchi

**HIGHLIGHT:** We consider the problem of separating a particular sound source from a single-channel mixture, based on only a short sample of the target source (from the same recording).

102, **TITLE:** Low Latency Online Blind Source Separation Based on Joint Optimization with Blind Dereverberation

<https://doi.org/10.1109/ICASSP39728.2021.9413700>

**AUTHORS:** T. Ueda, T. Nakatani, R. Ikeshita, K. Kinoshita, S. Araki and S. Makino

**HIGHLIGHT:** This paper proposes a method to solve this problem by integrating BSS with Weighted Prediction Error (WPE) based dereverberation.

103, **TITLE:** Autoregressive Fast Multichannel Nonnegative Matrix Factorization For Joint Blind Source Separation And Dereverberation

<https://doi.org/10.1109/ICASSP39728.2021.9414857>

**AUTHORS:** K. Sekiguchi, Y. Bando, A. A. Nugraha, M. Fontaine and K. Yoshii

**HIGHLIGHT:** This paper describes a joint blind source separation and dereverberation method that works adaptively and efficiently in a reverberant noisy environment.

104, **TITLE:** Phase Recovery with Bregman Divergences for Audio Source Separation

<https://doi.org/10.1109/ICASSP39728.2021.9413717>

**AUTHORS:** P. Magron, P. -H. Vial, T. Oberlin and C. F?votte

**HIGHLIGHT:** In this paper, we propose to reformulate phase recovery in audio source separation as a minimization problem involving Bregman divergences.

105, **TITLE:** Adversarial Attacks on Audio Source Separation

<https://doi.org/10.1109/ICASSP39728.2021.9414844>

**AUTHORS:** N. Takahashi, S. Inoue and Y. Mitsufuji

**HIGHLIGHT:** In this work, we reformulate various adversarial attack methods for the audio source separation problem and intensively investigate them under different attack conditions and target models.

106, **TITLE:** Maximum a Posteriori Estimator for Convolutional Sound Source Separation with Sub-Source Based NTF Model and the Localization Probabilistic Prior on the Mixing Matrix

<https://doi.org/10.1109/ICASSP39728.2021.9413863>

**AUTHORS:** M. Fras and K. Kowalczyk

**HIGHLIGHT:** In this paper we present a method for the separation of sound source signals recorded using multiple microphones in a reverberant room.

107, **TITLE:** Unified Gradient Reweighting for Model Biasing with Applications to Source Separation

<https://doi.org/10.1109/ICASSP39728.2021.9414071>

**AUTHORS:** E. Tzinis, D. Bralios and P. Smaragdis

**HIGHLIGHT:** In this paper, we propose a simple, unified gradient reweighting scheme, with a lightweight modification to bias the learning process of a model and steer it towards a certain distribution of results.

108, **TITLE:** Melon Playlist Dataset: A Public Dataset for Audio-Based Playlist Generation and Music Tagging

<https://doi.org/10.1109/ICASSP39728.2021.9413552>

**AUTHORS:** A. Ferraro et al.

**HIGHLIGHT:** We present Melon Playlist Dataset, a public dataset of mel-spectrograms for 649,091 tracks and 148,826 associated playlists annotated by 30,652 different tags.

109, **TITLE:** Investigating the Efficacy of Music Version Retrieval Systems for Setlist Identification

<https://doi.org/10.1109/ICASSP39728.2021.9414603>

AUTHORS: F. Yesiler, E. Molina, J. Serr? and E. G?mez

HIGHLIGHT: In this paper, we propose an end-to-end workflow that identifies relevant metadata and timestamps of live music performances using a version identification system.

110, TITLE: Instrument Classification of Solo Sheet Music Images

<https://doi.org/10.1109/ICASSP39728.2021.9413732>

AUTHORS: K. Ji, D. Yang and T. Tsai

HIGHLIGHT: In this work, we train AWD-LSTM, GPT-2, and RoBERTa models on solo sheet music images from IMSLP for eight different instruments.

111, TITLE: Bytecover: Cover Song Identification Via Multi-Loss Training

<https://doi.org/10.1109/ICASSP39728.2021.9414128>

AUTHORS: X. Du, Z. Yu, B. Zhu, X. Chen and Z. Ma

HIGHLIGHT: We present in this paper ByteCover, which is a new feature learning method for cover song identification (CSI).

112, TITLE: Multi-Task Self-Supervised Pre-Training for Music Classification

<https://doi.org/10.1109/ICASSP39728.2021.9414405>

AUTHORS: H. -H. Wu et al.

HIGHLIGHT: In this paper, we apply self-supervised and multi-task learning methods for pre-training music encoders, and explore various design choices including encoder architectures, weighting mechanisms to combine losses from multiple tasks, and worker selections of pretext tasks.

113, TITLE: Towards Explaining Expressive Qualities in Piano Recordings: Transfer of Explanatory Features Via Acoustic Domain Adaptation

<https://doi.org/10.1109/ICASSP39728.2021.9413638>

AUTHORS: S. Chowdhury and G. Widmer

HIGHLIGHT: In this work, we show that by utilising unsupervised domain adaptation together with receptive-field regularised deep neural networks, it is possible to significantly improve generalisation to this domain.

114, TITLE: Supervised Chorus Detection for Popular Music Using Convolutional Neural Network and Multi-Task Learning

<https://doi.org/10.1109/ICASSP39728.2021.9413773>

AUTHORS: J. -C. Wang, J. B. L. Smith, J. Chen, X. Song and Y. Wang

HIGHLIGHT: This paper presents a novel supervised approach to detecting the chorus segments in popular music.

115, TITLE: Structure-Aware Audio-to-Score Alignment Using Progressively Dilated Convolutional Neural Networks

<https://doi.org/10.1109/ICASSP39728.2021.9414049>

AUTHORS: R. Agrawal, D. Wolff and S. Dixon

HIGHLIGHT: We present a novel method to detect such differences between the score and performance for a given piece of music using progressively dilated convolutional neural networks.

116, TITLE: Language-Sensitive Music Emotion Recognition Models: are We Really There Yet?

<https://doi.org/10.1109/ICASSP39728.2021.9413721>

AUTHORS: J. S. G?mez-Ca??n, E. Cano, A. G. Pandrea, P. Herrera and E. G?mez

HIGHLIGHT: This paper presents additional investigation on our approach, which reveals that: (1) performing pretraining with speech in a mixture of languages yields similar results than for specific languages - the pretraining phase appears not to exploit particular language features, (2) the music in Mandarin dataset consistently results in poor classification performance - we found low agreement in annotations, and (3) novel methodologies for representation learning (Contrastive Predictive Coding) may exploit features from both languages (i.e., pretraining on a mixture of languages) and improve classification of music emotions in both languages.

117, TITLE: Leveraging the Structure of Musical Preference in Content-Aware Music Recommendation

<https://doi.org/10.1109/ICASSP39728.2021.9414194>

AUTHORS: P. Magron and C. F?votte

HIGHLIGHT: In this work, we propose instead to leverage a model of musical preference which originates from the field of music psychology.

118, TITLE: Low Resource Audio-To-Lyrics Alignment from Polyphonic Music Recordings

<https://doi.org/10.1109/ICASSP39728.2021.9414395>

AUTHORS: E. Demirel, S. Ahlb?ck and S. Dixon

**HIGHLIGHT:** In this study, we present a novel method that performs audio-to-lyrics alignment with a low memory consumption footprint regardless of the duration of the music recording.

119, **TITLE:** Multimodal Metric Learning for Tag-Based Music Retrieval

<https://doi.org/10.1109/ICASSP39728.2021.9413514>

**AUTHORS:** M. Won, S. Oramas, O. Nieto, F. Gouyon and X. Serra

**HIGHLIGHT:** In this paper, we investigate three ideas to successfully introduce multimodal metric learning for tag-based music retrieval: elaborate triplet sampling, acoustic and cultural music information, and domain-specific word embeddings.

120, **TITLE:** Learning Contextual Tag Embeddings for Cross-Modal Alignment of Audio and Tags

<https://doi.org/10.1109/ICASSP39728.2021.9414638>

**AUTHORS:** X. Favory, K. Drossos, T. Virtanen and X. Serra

**HIGHLIGHT:** In this work we propose a method for learning audio representations using an audio autoencoder (AAE), a general word embeddings model (WEM), and a multi-head self-attention (MHA) mechanism.

121, **TITLE:** Efficient End-to-End Audio Embeddings Generation for Audio Classification on Target Applications

<https://doi.org/10.1109/ICASSP39728.2021.9414229>

**AUTHORS:** P. Lopez-Meyer, J. A. del Hoyo Ontiveros, H. Lu and G. Stemmer

**HIGHLIGHT:** We describe a general-purpose end-to-end audio embeddings generator that can be easily adapted to various acoustic scene and event classification applications.

122, **TITLE:** Text-to-Audio Grounding: Building Correspondence Between Captions and Sound Events

<https://doi.org/10.1109/ICASSP39728.2021.9414834>

**AUTHORS:** X. Xu, H. Dinkel, M. Wu and K. Yu

**HIGHLIGHT:** Based on such, we propose the text-to-audio grounding (TAG) task, which interactively considers the relationship between audio processing and language understanding.

123, **TITLE:** Multi-View Audio And Music Classification

<https://doi.org/10.1109/ICASSP39728.2021.9414551>

**AUTHORS:** H. Phan et al.

**HIGHLIGHT:** We propose in this work a multi-view learning approach for audio and music classification.

124, **TITLE:** Audio-Visual Event Recognition Through the Lens of Adversary

<https://doi.org/10.1109/ICASSP39728.2021.9415065>

**AUTHORS:** J. B. Li, K. Ma, S. Qu, P. -Y. Huang and F. Metze

**HIGHLIGHT:** This work aims to study several key questions related to multimodal learning through the lens of adversarial noises: 1) The trade-off between early/middle/late fusion affecting its robustness and accuracy 2) How does different frequency/time domain features contribute to the robustness?

125, **TITLE:** DCASENET: An Integrated Pretrained Deep Neural Network for Detecting and Classifying Acoustic Scenes and Events

<https://doi.org/10.1109/ICASSP39728.2021.9414406>

**AUTHORS:** J. -w. Jung, H. -j. Shim, J. -h. Kim and H. -J. Yu

**HIGHLIGHT:** We propose three architectures of deep neural networks that are integrated to simultaneously perform acoustic scene classification, audio tagging, and sound event detection.

126, **TITLE:** A Curated Dataset of Urban Scenes for Audio-Visual Scene Analysis

<https://doi.org/10.1109/ICASSP39728.2021.9415085>

**AUTHORS:** S. Wang, A. Mesaros, T. Heittola and T. Virtanen

**HIGHLIGHT:** This paper introduces a curated dataset of urban scenes for audio-visual scene analysis which consists of carefully selected and recorded material.

127, **TITLE:** Improving Sound Event Detection Metrics: Insights from DCASE 2020

<https://doi.org/10.1109/ICASSP39728.2021.9414711>

**AUTHORS:** G. Ferroni et al.

**HIGHLIGHT:** This paper compares conventional event-based and segment-based criteria against the Polyphonic Sound Detection Score (PSDS)'s intersection-based criterion, over a selection of systems from DCASE 2020 Challenge Task 4.

128, **TITLE:** Artificially Synthesising Data for Audio Classification and Segmentation to Improve Speech and Music Detection in Radio Broadcast

<https://doi.org/10.1109/ICASSP39728.2021.9413597>

AUTHORS: S. Venkatesh et al.

HIGHLIGHT: In this study, we present a novel procedure that artificially synthesises data that resembles radio signals.

129, TITLE: LSSED: A Large-Scale Dataset and Benchmark for Speech Emotion Recognition

<https://doi.org/10.1109/ICASSP39728.2021.9414542>

AUTHORS: W. Fan, X. Xu, X. Xing, W. Chen and D. Huang

HIGHLIGHT: In this paper, we present LSSED, a challenging large-scale english speech emotion dataset, which has data collected from 820 subjects to simulate real- world distribution.

130, TITLE: Enhancing Audio Augmentation Methods with Consistency Learning

<https://doi.org/10.1109/ICASSP39728.2021.9414316>

AUTHORS: T. Iqbal, K. Helwani, A. Krishnaswamy and W. Wang

HIGHLIGHT: This paper investigates the use of training objectives that explicitly impose this consistency constraint, and how it can impact downstream audio classification tasks.

131, TITLE: Fast Threshold Optimization for Multi-Label Audio Tagging Using Surrogate Gradient Learning

<https://doi.org/10.1109/ICASSP39728.2021.9414091>

AUTHORS: T. Pellegrini and T. Masquelier

HIGHLIGHT: In this work, we consider having at disposal a trained classifier and we seek to automatically optimize the decision thresholds according to a performance metric of interest, in our case F-measure (micro-F1).

132, TITLE: Towards Efficient Models for Real-Time Deep Noise Suppression

<https://doi.org/10.1109/ICASSP39728.2021.9413580>

AUTHORS: S. Braun, H. Gamper, C. K. A. Reddy and I. Tashev

HIGHLIGHT: In this work, we investigate reasonably small recurrent and convolutional-recurrent network architectures for speech enhancement, trained on a large dataset considering also reverberation.

133, TITLE: Teacher-Student Learning for Low-Latency Online Speech Enhancement Using Wave-U-Net

<https://doi.org/10.1109/ICASSP39728.2021.9414280>

AUTHORS: S. Nakaoka, L. Li, S. Inoue and S. Makino

HIGHLIGHT: In this paper, we propose a low-latency online extension of wave-U-net for single-channel speech enhancement, which utilizes teacher-student learning to reduce the system latency while keeping the enhancement performance high.

134, TITLE: Learning Disentangled Feature Representations for Speech Enhancement Via Adversarial Training

<https://doi.org/10.1109/ICASSP39728.2021.9413512>

AUTHORS: N. Hou, C. Xu, E. S. Chng and H. Li

HIGHLIGHT: To address such mismatch, we propose to learn noise-agnostic feature representations by disentanglement learning, which removes the unspecified noise factor, while keeping the specified factors of variation associated with the clean speech.

135, TITLE: Speech Enhancement Autoencoder with Hierarchical Latent Structure

<https://doi.org/10.1109/ICASSP39728.2021.9414695>

AUTHORS: K. Oostermeijer, J. Du, Q. Wang and C. -H. Lee

HIGHLIGHT: A new hierarchical convolutional neural network-based autoencoder architecture called SEHAE (Speech Enhancement Hierarchical AutoEncoder) is introduced, in which the latent representation is decomposed into several parts that correspond to different scales.

136, TITLE: Variational Autoencoder for Speech Enhancement with a Noise-Aware Encoder

<https://doi.org/10.1109/ICASSP39728.2021.9414060>

AUTHORS: H. Fang, G. Carbajal, S. Wermter and T. Gerkmann

HIGHLIGHT: To increase the robustness of the VAE, we propose to include noise information in the training phase by using a noise-aware encoder trained on noisy-clean speech pairs.

137, TITLE: Guided Variational Autoencoder for Speech Enhancement with a Supervised Classifier

<https://doi.org/10.1109/ICASSP39728.2021.9414363>

AUTHORS: G. Carbajal, J. Richter and T. Gerkmann

HIGHLIGHT: In this paper, we propose to guide the variational autoencoder with a supervised classifier separately trained on noisy speech.

138, TITLE: An Extension of Sparse Audio Declipper to Multiple Measurement Vectors

<https://doi.org/10.1109/ICASSP39728.2021.9414519>

AUTHORS: S. Emura and N. Harada

HIGHLIGHT: This paper proposes formulating declipping as a constrained multiple measurement vector (MMV) optimization problem that has a  $\ell_{2,0}$  group norm as its cost function for further improving the state-of-the-art declipping method SParse Audio DEclipper (SPADE).

139, TITLE: Real-Time Speech Frequency Bandwidth Extension

<https://doi.org/10.1109/ICASSP39728.2021.9413439>

AUTHORS: Y. Li, M. Tagliasacchi, O. Rybakov, V. Ungureanu and D. Roblek

HIGHLIGHT: In this paper we propose a lightweight model for frequency bandwidth extension of speech signals, increasing the sampling frequency from 8kHz to 16kHz while restoring the high frequency content to a level almost indistinguishable from the 16kHz ground truth.

140, TITLE: Bandwidth Extension is All You Need

<https://doi.org/10.1109/ICASSP39728.2021.9413575>

AUTHORS: J. Su, Y. Wang, A. Finkelstein and Z. Jin

HIGHLIGHT: This paper proposes a new bandwidth extension (BWE) method that expands 8-16kHz speech signals to 48kHz.

141, TITLE: Audio Dequantization Using (Co)Sparse (Non)Convex Methods

<https://doi.org/10.1109/ICASSP39728.2021.9414637>

AUTHORS: P. Z?vi?ka, P. Rajmic and O. Mokr?

HIGHLIGHT: It reviews the state-of-the-art sparsity-based approaches and proposes several new methods.

142, TITLE: Source-Aware Neural Speech Coding for Noisy Speech Compression

<https://doi.org/10.1109/ICASSP39728.2021.9413678>

AUTHORS: H. Yang, K. Zhen, S. Beack and M. Kim

HIGHLIGHT: This paper introduces a novel neural network-based speech coding system that can process noisy speech effectively.

143, TITLE: Enhancing into the Codec: Noise Robust Speech Coding with Vector-Quantized Autoencoders

<https://doi.org/10.1109/ICASSP39728.2021.9414605>

AUTHORS: J. Casebeer, V. Vale, U. Isik, J. -M. Valin, R. Giri and A. Krishnaswamy

HIGHLIGHT: Based on VQ-VAE autoencoders with WaveRNN decoders, we develop compressor-enhancer encoders and accompanying decoders, and show that they operate well in noisy conditions.

144, TITLE: Speech Enhancement with Mixture of Deep Experts with Clean Clustering Pre-Training

<https://doi.org/10.1109/ICASSP39728.2021.9414122>

AUTHORS: S. E. Chazan, J. Goldberger and S. Gannot

HIGHLIGHT: In this study we present a mixture of deep experts (MoDE) neural-network architecture for single microphone speech enhancement.

145, TITLE: A Novel NMF-HMM Speech Enhancement Algorithm Based on Poisson Mixture Model

<https://doi.org/10.1109/ICASSP39728.2021.9414620>

AUTHORS: Y. Xiang, L. Shi, J. L. H?jvang, M. H?jfeldt Rasmussen and M. G. Christensen

HIGHLIGHT: In this paper, we propose a novel non-negative matrix factorization (NMF) and hidden Markov model (NMF-HMM) based speech enhancement algorithm, which employs a Poisson mixture model (PMM).

146, TITLE: Phoneme-Based Distribution Regularization for Speech Enhancement

<https://doi.org/10.1109/ICASSP39728.2021.9414761>

AUTHORS: Y. Liu, X. Peng, Z. Xiong and Y. Lu

HIGHLIGHT: In this paper, we aim to bridge this gap by extracting phoneme identities to help speech enhancement.

147, TITLE: Compressed Representation of Cepstral Coefficients via Recurrent Neural Networks for Informed Speech Enhancement

<https://doi.org/10.1109/ICASSP39728.2021.9414536>

AUTHORS: C. Chermaz, D. Leuchtman, S. Tanner and R. Wattenhofer

HIGHLIGHT: We investigate a hybrid strategy made of signal processing and RNN (Recurrent Neural Networks) to calculate and compress cepstral coefficients: these are descriptors of the speech signal, which can be embedded in the signal itself and used at the receiver's end to perform an Informed Speech Enhancement.



- 148, TITLE: Optimizing Short-Time Fourier Transform Parameters via Gradient Descent  
<https://doi.org/10.1109/ICASSP39728.2021.9413704>  
AUTHORS: A. Zhao, K. Subramani and P. Smaragdis  
HIGHLIGHT: In this paper we show an approach that allows us to obtain a gradient for STFT parameters with respect to arbitrary cost functions, and thus enable the ability to employ gradient descent optimization of quantities like the STFT window length, or the STFT hop size.
- 149, TITLE: Iterative Geometry Calibration from Distance Estimates for Wireless Acoustic Sensor Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9413831>  
AUTHORS: T. Gburrek, J. Schmalenstroeeer and R. Haeb-Umbach  
HIGHLIGHT: In this paper we present an approach to geometry calibration in wireless acoustic sensor networks, whose nodes are assumed to be equipped with a compact microphone array.
- 150, TITLE: On the Design of Square Differential Microphone Arrays with a Multistage Structure  
<https://doi.org/10.1109/ICASSP39728.2021.9413759>  
AUTHORS: X. Zhao, G. Huang, J. Benesty, J. Chen and I. Cohen  
HIGHLIGHT: It presents a multistage approach, which first divides an SDMA composed of  $M_2$  microphones into  $(M - 1)2$  subarrays with each subarray being a  $2 \times 2$  square array formed by four adjacent microphones.
- 151, TITLE: Arrays of First-Order Steerable Differential Microphones  
<https://doi.org/10.1109/ICASSP39728.2021.9413476>  
AUTHORS: F. Borra, A. Bernardini, I. Bertuletti, F. Antonacci and A. Sarti  
HIGHLIGHT: In this paper, we consider arbitrarily shaped planar arrays of DMA units.
- 152, TITLE: Planar Array Geometry Optimization for Region Sound Acquisition  
<https://doi.org/10.1109/ICASSP39728.2021.9414446>  
AUTHORS: X. Chen, C. Pan, J. Chen and J. Benesty  
HIGHLIGHT: This paper studies the problem of geometry optimization for planar arrays and it develops a genetic optimization algorithm that can optimize the positions of the sensors, thereby maximizing the directivity factor (DF) with a constrained level of white noise gain (WNG) given the number of microphones, the region in which they should be placed, and the interested range of steering.
- 153, TITLE: Estimation of Microphone Clusters in Acoustic Sensor Networks Using Unsupervised Federated Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9414186>  
AUTHORS: A. Nelus, R. Glitza and R. Martin  
HIGHLIGHT: In this paper we present a privacy-aware method for estimating source-dominated microphone clusters in the context of acoustic sensor networks (ASNs).
- 154, TITLE: Misalignment Recognition in Acoustic Sensor Networks Using a Semi-Supervised Source Estimation Method and Markov Random Fields  
<https://doi.org/10.1109/ICASSP39728.2021.9413765>  
AUTHORS: G. F. Miller, A. Brendel, W. Kellermann and S. Gannot  
HIGHLIGHT: In this paper, we consider the problem of acoustic source localization by acoustic sensor networks (ASNs) using a promising, learning-based technique that adapts to the acoustic environment.
- 155, TITLE: Rotation-Robust Beamforming Based on Sound Field Interpolation with Regularly Circular Microphone Array  
<https://doi.org/10.1109/ICASSP39728.2021.9414196>  
AUTHORS: Y. Wakabayashi, K. Yamaoka and N. Ono  
HIGHLIGHT: In this paper, we present a novel framework of beamforming robust for a microphone array rotation.
- 156, TITLE: Sparse Recovery Beamforming and Upscaling in the Ray Space  
<https://doi.org/10.1109/ICASSP39728.2021.9414268>  
AUTHORS: S. Yu, C. Jin, F. Antonacci and A. Sarti  
HIGHLIGHT: In this work, we explore a method to upscale an array beyond the limits imposed by the inter-microphone distances associated with the array and the concomitant spatial aliasing.
- 157, TITLE: Combined Differential Beamforming With Uniform Linear Microphone Arrays  
<https://doi.org/10.1109/ICASSP39728.2021.9414189>  
AUTHORS: G. Huang, Y. Wang, J. Benesty, I. Cohen and J. Chen  
HIGHLIGHT: It presents a method for the design of differential beamformers with uniform linear arrays.

- 158, TITLE: Polynomial Matrix Eigenvalue Decomposition of Spherical Harmonics for Speech Enhancement  
<https://doi.org/10.1109/ICASSP39728.2021.9414011>  
AUTHORS: V. W. Neo, C. Evers and P. A. Naylor  
HIGHLIGHT: We propose a PEVD algorithm that uses only the lower dimension eigenbeams for speech enhancement at a significantly lower computation cost.
- 159, TITLE: A Parametric Unconstrained Binaural Beamformer Based Noise Reduction and Spatial Cue Preservation for Hearing-Assistive Devices  
<https://doi.org/10.1109/ICASSP39728.2021.9414210>  
AUTHORS: J. Zhang  
HIGHLIGHT: In this work, we propose a parametric unconstrained binaural (PUB) beamformer, which can achieve a trade-off between noise reduction and binaural cue preservation.
- 160, TITLE: A Simplified Wiener Beamformer Based on Covariance Matrix Modelling  
<https://doi.org/10.1109/ICASSP39728.2021.9414719>  
AUTHORS: F. Zhang, C. Pan, J. Benesty and J. Chen  
HIGHLIGHT: To deal with this problem, we propose a general method by parametric modeling the covariance matrices of speech and noise, which leads to a simplified Wiener beamformer.
- 161, TITLE: Control Architecture of the Double-Cross-Correlation Processor for Sampling-Rate-Offset Estimation in Acoustic Sensor Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9413768>  
AUTHORS: A. Chinaev, S. Wienand and G. Enzner  
HIGHLIGHT: This paper converts the mechanism of offline multi-stage processing into a continuous feedback-control loop comprising a controlled ASRC unit followed by an online implementation of DXCP-based SRO estimation.
- 162, TITLE: Deficient Basis Estimation of Noise Spatial Covariance Matrix for Rank-Constrained Spatial Covariance Matrix Estimation Method in Blind Speech Extraction  
<https://doi.org/10.1109/ICASSP39728.2021.9414479>  
AUTHORS: Y. Kondo, Y. Kubo, N. Takamune, D. Kitamura and H. Saruwatari  
HIGHLIGHT: In this paper, we proposed a new algorithmic extension of RCSCME.
- 163, TITLE: Reducing Modal Error Propagation through Correcting Mismatched Microphone Gains Using Rapid  
<https://doi.org/10.1109/ICASSP39728.2021.9415030>  
AUTHORS: N. Akbar, G. Dickins, M. R. P. Thomas, P. Samarasinghe and T. Abhayapala  
HIGHLIGHT: A method for reducing the error propagation in modes by correcting the mismatched microphone gains is proposed, where RANdom Perturbations for Diffuse-field (RAPID) is used to design filters for correcting the mismatch.
- 164, TITLE: Evaluation and Comparison of Three Source Direction-of-Arrival Estimators Using Relative Harmonic Coefficients  
<https://doi.org/10.1109/ICASSP39728.2021.9414481>  
AUTHORS: Y. Hu, P. N. Samarasinghe, S. Gannot and T. D. Abhayapala  
HIGHLIGHT: This paper presents a compact evaluation and comparison between two existing RHC based DOA estimators: (i) a method using a full grid search over the two-dimensional (2-D) directional space, (ii) a decoupled estimator which uses one-dimensional (1-D) search to separately localize the source's elevation and azimuth.
- 165, TITLE: Network-Aware Optimal Microphone Channel Selection in Wireless Acoustic Sensor Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414528>  
AUTHORS: M. Gunther, H. Afifi, A. Brendel, H. Karl and W. Kellermann  
HIGHLIGHT: To address the vital problem of selecting the most useful microphones in wireless acoustic sensor networks, this paper proposes a novel, general-purpose approach that accounts for both acoustic and network aspects and remains application-agnostic for broad applicability.
- 166, TITLE: Supervised Direct-Path Relative Transfer Function Learning for Binaural Sound Source Localization  
<https://doi.org/10.1109/ICASSP39728.2021.9413923>  
AUTHORS: B. Yang, X. Li and H. Liu  
HIGHLIGHT: This paper proposes a supervised DP-RTF learning method with deep neural networks for robust binaural sound source localization.
- 167, TITLE: Cross-Modal Spectrum Transformation Network for Acoustic Scene Classification

<https://doi.org/10.1109/ICASSP39728.2021.9414779>

AUTHORS: Y. Liu, A. Neophytou, S. Sengupta and E. Sommerlade

HIGHLIGHT: To address this issue, we introduce an acoustic spectrum transformation network where traditional log-mel spectrums are transformed into imagined visual features (IVF).

168, TITLE: Domestic Activities Clustering From Audio Recordings Using Convolutional Capsule Autoencoder Network

<https://doi.org/10.1109/ICASSP39728.2021.9414643>

AUTHORS: Z. Lin et al.

HIGHLIGHT: In this study, we propose a method for domestic activities clustering using a convolutional capsule autoencoder network (CCAN).

169, TITLE: Sound Event Detection and Separation: A Benchmark on Desed Synthetic Soundscapes

<https://doi.org/10.1109/ICASSP39728.2021.9414789>

AUTHORS: N. Turpault et al.

HIGHLIGHT: We propose a benchmark of state-of-the-art sound event detection systems (SED).

170, TITLE: A Two-Stage Approach to Device-Robust Acoustic Scene Classification

<https://doi.org/10.1109/ICASSP39728.2021.9414835>

AUTHORS: H. Hu et al.

HIGHLIGHT: To improve device robustness, a highly desirable key feature of a competitive data-driven acoustic scene classification (ASC) system, a novel two-stage system based on fully convolutional neural networks (CNNs) is proposed.

171, TITLE: Subspectral Normalization for Neural Audio Data Processing

<https://doi.org/10.1109/ICASSP39728.2021.9413522>

AUTHORS: S. Chang, H. Park, J. Cho, H. Park, S. Yun and K. Hwang

HIGHLIGHT: In this work, we introduce SubSpectral Normalization (SSN), which splits the input frequency dimension into several groups (sub-bands) and performs a different normalization for each group.

172, TITLE: Slow-Fast Auditory Streams for Audio Recognition

<https://doi.org/10.1109/ICASSP39728.2021.9413376>

AUTHORS: E. Kazakos, A. Nagrani, A. Zisserman and D. Damen

HIGHLIGHT: We propose a two-stream convolutional network for audio recognition, that operates on time-frequency spectrogram inputs.

173, TITLE: Impact of Sound Duration and Inactive Frames on Sound Event Detection Performance

<https://doi.org/10.1109/ICASSP39728.2021.9414949>

AUTHORS: K. Imoto, S. Mishima, Y. Arai and R. Kondo

HIGHLIGHT: In this paper, we investigate the impact of sound duration and inactive frames on SED performance by introducing four loss functions, such as simple reweighting loss, inverse frequency loss, asymmetric focal loss, and focal batch Tversky loss.

174, TITLE: A New DCASE 2017 Rare Sound Event Detection Benchmark Under Equal Training Data: CRNN With Multi-Width Kernels

<https://doi.org/10.1109/ICASSP39728.2021.9414254>

AUTHORS: J. Baumann, P. Meyer, T. Lohrenz, A. Roy, M. Papendieck and T. Fingscheidt

HIGHLIGHT: In this work, we propose a new CRNN model for rare SED.

175, TITLE: Room Adaptive Conditioning Method for Sound Event Classification in Reverberant Environments

<https://doi.org/10.1109/ICASSP39728.2021.9413929>

AUTHORS: J. Lee, D. Lee, H. -S. Choi and K. Lee

HIGHLIGHT: To alleviate this problem, we propose a conditioning method that provides room impulse response (RIR) information to help the network become less sensitive to environmental information and focus on classifying the desired sound.

176, TITLE: Sound Event Detection Based on Curriculum Learning Considering Learning Difficulty of Events

<https://doi.org/10.1109/ICASSP39728.2021.9414184>

AUTHORS: N. Tonami, K. Imoto, Y. Okamoto, T. Fukumori and Y. Yamashita

HIGHLIGHT: To utilize the curriculum learning, we propose a new objective function for SED, wherein the events are trained from easy-to difficult-to-train events.

177, TITLE: Sound Event Detection in Urban Audio with Single and Multi-Rate Pcen

<https://doi.org/10.1109/ICASSP39728.2021.9414697>

AUTHORS: C. Ick and B. McFee

HIGHLIGHT: In this article, we experiment using PCEN spectrograms as an alternative method for SED in urban audio using the UrbanSED dataset, demonstrating per-class improvements based on parameter configuration.

178, TITLE: An Improved Event-Independent Network for Polyphonic Sound Event Localization and Detection

<https://doi.org/10.1109/ICASSP39728.2021.9413473>

AUTHORS: Y. Cao, T. Iqbal, Q. Kong, F. An, W. Wang and M. D. Plumbley

HIGHLIGHT: Two open problems are addressed in this paper. Firstly, to detect overlapping sound events of the same type but with different DoAs, we propose to use a trackwise output format and solve the accompanying track permutation problem with permutation-invariant training. Multi-head self-attention is further used to separate tracks. Secondly, a previous finding is that, by using hard parameter-sharing, SELD suffers from a performance loss compared with learning the subtasks separately.

179, TITLE: Lightweight and Interpretable Neural Modeling of an Audio Distortion Effect Using Hyperconditioned Differentiable Biquads

<https://doi.org/10.1109/ICASSP39728.2021.9413996>

AUTHORS: S. Nercessian, A. Sarroff and K. J. Werner

HIGHLIGHT: In this work, we propose using differentiable cascaded biquads to model an audio distortion effect.

180, TITLE: Attacking and Defending Behind A Psychoacoustics-Based Captcha

<https://doi.org/10.1109/ICASSP39728.2021.9414135>

AUTHORS: C. -H. Huang, P. -H. Wu, Y. -W. Liu and S. -H. Wu

HIGHLIGHT: This paper proposes a novel audio CAPTCHA system that requires a user to respond immediately after hearing a short and easy-to-remember cue in its mixture with background music.

181, TITLE: Double-DCCCAE: Estimation of Body Gestures From Speech Waveform

<https://doi.org/10.1109/ICASSP39728.2021.9414660>

AUTHORS: J. Lu, T. Liu, S. Xu and H. Shimodaira

HIGHLIGHT: This paper presents an approach for body-motion estimation from audio-speech waveform, where context information in both input and output streams is taken in to account without using recurrent models.

182, TITLE: Investigating Local and Global Information for Automated Audio Captioning with Transfer Learning

<https://doi.org/10.1109/ICASSP39728.2021.9413982>

AUTHORS: X. Xu, H. Dinkel, M. Wu, Z. Xie and K. Yu

HIGHLIGHT: This paper first proposes a topic model for audio descriptions, comprehensively analyzing the hierarchical audio topics that are commonly covered. We then explore a transfer learning scheme to access local and global information.

183, TITLE: Unidirectional Memory-Self-Attention Transducer for Online Speech Recognition

<https://doi.org/10.1109/ICASSP39728.2021.9413932>

AUTHORS: J. Luo, J. Wang, N. Cheng and J. Xiao

HIGHLIGHT: In this paper, we propose Memory-Self-Attention (MSA), which adds history information into the Restricted-Self-Attention unit.

184, TITLE: Accdoa: Activity-Coupled Cartesian Direction of Arrival Representation for Sound Event Localization And Detection

<https://doi.org/10.1109/ICASSP39728.2021.9413609>

AUTHORS: K. Shimada, Y. Koyama, N. Takahashi, S. Takahashi and Y. Mitsufuji

HIGHLIGHT: To address these problems, we propose an activity-coupled Cartesian DOA (ACCDOA) representation, which assigns a sound event activity to the length of a corresponding Cartesian DOA vector.

185, TITLE: Seen and Unseen Emotional Style Transfer for Voice Conversion with A New Emotional Speech Dataset

<https://doi.org/10.1109/ICASSP39728.2021.9413391>

AUTHORS: K. Zhou, B. Sisman, R. Liu and H. Li

HIGHLIGHT: In this paper, we propose a novel framework based on variational auto-encoding Wasserstein generative adversarial network (VAW-GAN), which makes use of a pre-trained speech emotion recognition (SER) model to transfer emotional style during training and at run-time inference.

186, TITLE: U-Convolution Based Residual Echo Suppression with Multiple Encoders

<https://doi.org/10.1109/ICASSP39728.2021.9414753>

AUTHORS: E. Kim, J. -J. Jeon and H. Seo

**HIGHLIGHT:** In this paper, we propose an efficient end-to-end neural network that can estimate near-end speech using a U-convolution block by exploiting various signals to achieve residual echo suppression (RES).

187, **TITLE:** A Multi-Channel Temporal Attention Convolutional Neural Network Model for Environmental Sound Classification

<https://doi.org/10.1109/ICASSP39728.2021.9413498>

**AUTHORS:** Y. Wang, C. Feng and D. V. Anderson

**HIGHLIGHT:** In this paper, we propose an effective convolutional neural network structure with a multichannel temporal attention (MCTA) block, which applies a temporal attention mechanism within each channel of the embedded features to extract channel-wise relevant temporal information.

188, **TITLE:** A General Network Architecture for Sound Event Localization and Detection Using Transfer Learning and Recurrent Neural Network

<https://doi.org/10.1109/ICASSP39728.2021.9414602>

**AUTHORS:** T. N. T. Nguyen et al.

**HIGHLIGHT:** We propose a general network architecture for SELD in which the SELD network comprises sub-networks that are pre-trained to solve SED and DOA estimation independently, and a recurrent layer that combines the SED and DOA estimation outputs into SELD outputs.

189, **TITLE:** Robust Recursive Least M-Estimate Adaptive Filter for the Identification of Low-Rank Acoustic Systems

<https://doi.org/10.1109/ICASSP39728.2021.9413983>

**AUTHORS:** H. He, J. Chen, J. Benesty and Y. Yu

**HIGHLIGHT:** To identify acoustic systems (which are low-rank in nature) in non-Gaussian and Gaussian noise, a robust recursive least M-estimate adaptive filtering algorithm is developed in this paper by applying the nearest Kronecker product to decompose the acoustic impulse response.

190, **TITLE:** Noise-Robust Adaptation Control for Supervised Acoustic System Identification Exploiting a Noise Dictionary

<https://doi.org/10.1109/ICASSP39728.2021.9414180>

**AUTHORS:** T. Haubner, A. Brendel, M. Elminshawy and W. Kellermann

**HIGHLIGHT:** We present a noise-robust adaptation control strategy for block-online supervised acoustic system identification by exploiting a noise dictionary.

191, **TITLE:** Interpolation of Irregularly Sampled Frequency Response Functions Using Convolutional Neural Networks

<https://doi.org/10.1109/ICASSP39728.2021.9413458>

**AUTHORS:** M. Acerbi, R. Malvermi, M. Pezzoli, F. Antonacci, A. Sarti and R. Corradi

**HIGHLIGHT:** In this paper we propose to use Convolutional Autoencoders (CA) for Frequency Response Function (FRF) interpolation from grids with different subsampling schemes.

192, **TITLE:** Effective Rank-Based Estimation of the Coherent-to-Diffuse Power Ratio

<https://doi.org/10.1109/ICASSP39728.2021.9413985>

**AUTHORS:** H. W. L?llmann, A. Brendel and W. Kellermann

**HIGHLIGHT:** A CDR estimator whose design is based on this premise is devised in this contribution.

193, **TITLE:** Room Impulse Response Interpolation from a Sparse Set of Measurements Using a Modal Architecture

<https://doi.org/10.1109/ICASSP39728.2021.9414399>

**AUTHORS:** O. Das, P. Calamia and S. V. Amengual Gari

**HIGHLIGHT:** In this paper, we propose a novel method for 2D interpolation of room modes from a sparse set of RIR measurements that are non-uniformly sampled within a space.

194, **TITLE:** Processing Pipelines for Efficient, Physically-Accurate Simulation of Microphone Array Signals in Dynamic Sound Scenes

<https://doi.org/10.1109/ICASSP39728.2021.9413354>

**AUTHORS:** A. H. Moore, R. R. Vos, P. A. Naylor and M. Brookes

**HIGHLIGHT:** A new approach, in which the filter kernels are obtained using principal component analysis from time-aligned impulse responses, is proposed.

195, **TITLE:** A Classifier for Improving Cause and Effect in SSVEP-based BCIs for Individuals with Complex Communication Disorders

<https://doi.org/10.1109/ICASSP39728.2021.9414075>

**AUTHORS:** H. Habibzadeh, O. Zhou, J. J. S. Norton, T. M. Vaughan and D. -S. Zois

**HIGHLIGHT:** We present CCACUSUM, a classifier for steady-state visual evoked potential (SSVEP)-based brain-computer interfaces (BCIs) that determines whether a user is attending to a flickering stimulus or is at rest.

196, **TITLE:** Saga: Sparse Adversarial Attack on EEG-Based Brain Computer Interface

<https://doi.org/10.1109/ICASSP39728.2021.9413507>

**AUTHORS:** B. Feng, Y. Wang and Y. Ding

**HIGHLIGHT:** In this paper, we conduct the first in-depth study on the robustness of EEG analytics under sparse perturbations and propose the first Sparse Adversarial eeG Attack, SAGA, to identify weakness of EEG analytics.

197, **TITLE:** Riemannian Geometry on Connectivity for Clinical BCI

<https://doi.org/10.1109/ICASSP39728.2021.9414790>

**AUTHORS:** M. -C. Corsi, F. Yger, S. Chevallier and C. No?s

**HIGHLIGHT:** To increase the accuracy of BCI systems, we propose an approach grounded on Riemannian geometry that extends this framework to functional connectivity measures.

198, **TITLE:** Decoding Music Attention from ?EEG Headphones?: A User-Friendly Auditory Brain-Computer Interface

<https://doi.org/10.1109/ICASSP39728.2021.9414492>

**AUTHORS:** W. W. An et al.

**HIGHLIGHT:** We propose a novel BCI system using music stimuli that relies on brain signals collected via Smartfones, an EEG recording device integrated into a pair of headphones.

199, **TITLE:** Mitigating Inter-Subject Brain Signal Variability FOR EEG-Based Driver Fatigue State Classification

<https://doi.org/10.1109/ICASSP39728.2021.9414613>

**AUTHORS:** S. Hwang, S. Park, D. Kim, J. Lee and H. Byun

**HIGHLIGHT:** In this paper, we propose a subject- independent EEG-based driver fatigue state (i.e., awake, tired, and drowsy) classification model that mitigates a performance gap between subjects.

200, **TITLE:** A Deep Spatio-Temporal Model for EEG-Based Imagined Speech Recognition

<https://doi.org/10.1109/ICASSP39728.2021.9413989>

**AUTHORS:** P. Kumar and E. Scheme

**HIGHLIGHT:** Consequently, in this work, we propose an imagined speech Brain-Computer-Interface (BCI) using Electroencephalogram (EEG) signals.

201, **TITLE:** Incorporating Uncertainty In Data Labeling Into Detection of Brain Interictal Epileptiform Discharges From EEG Using Weighted optimization

<https://doi.org/10.1109/ICASSP39728.2021.9414463>

**AUTHORS:** B. Abdi-Sargezeh, A. Valentin, G. Alarcon and S. Sanei

**HIGHLIGHT:** Here, we incorporate this probability in an IED detection system which combines spatial component analysis (SCA) with the IED probabilities referred to as SCA-IEDP-based method.

202, **TITLE:** Multi-Level Reversible Encryption for ECG Signals Using Compressive Sensing

<https://doi.org/10.1109/ICASSP39728.2021.9414983>

**AUTHORS:** M. Impi?, M. Yama? and J. Raitoharju

**HIGHLIGHT:** In this paper, we propose a compressive sensing based multi-level encryption to ECG signals to mask possible heartbeat anomalies from semi-authorized users, while preserving the beat structure for heart rate monitoring.

203, **TITLE:** Validating the Inspired Sinewave Technique to Measure Lung Heterogeneity Compared to Atelectasis & Over-Distended Volume in Computed Tomography Images

<https://doi.org/10.1109/ICASSP39728.2021.9413942>

**AUTHORS:** M. C. Tran et al.

**HIGHLIGHT:** Six anaesthetised pigs were studied after surfactant depletion by saline-lavage.

204, **TITLE:** A Patient-Invariant Model for Freezing of Gait Detection Aided by Wavelet Decomposition

<https://doi.org/10.1109/ICASSP39728.2021.9414780>

**AUTHORS:** N. Ahmed, S. Singhal, V. Sharma, S. Bhattacharya, A. Sinha and A. Ghose

**HIGHLIGHT:** In this paper, we present a method for online detection of FoG using a wearable motion sensor.

205, **TITLE:** Identification of Uterine Contractions by An Ensemble of Gaussian Processes

<https://doi.org/10.1109/ICASSP39728.2021.9414041>

**AUTHORS:** L. Yang, C. Heiselman, J. Gerald Quirk and P. M. Djuri

**HIGHLIGHT:** In this paper, we study contraction identification by processing noisy signals due to uterine activities.

206, **TITLE:** Arrhythmia Classification with Heartbeat-Aware Transformer  
<https://doi.org/10.1109/ICASSP39728.2021.9413938>  
**AUTHORS:** B. Wang, C. Liu, C. Hu, X. Liu and J. Cao  
**HIGHLIGHT:** In this paper, we proposed a novel neural network model which treats typical heartbeat classification task as 'Translation' problem.

207, **TITLE:** Multi-Level Group Testing with Application to One-Shot Pooled COVID-19 Tests  
<https://doi.org/10.1109/ICASSP39728.2021.9414574>  
**AUTHORS:** A. Cohen, N. Shlezinger, A. Solomon, Y. C. Eldar and M. M?dard  
**HIGHLIGHT:** In this work we study pooling-based COVID-19 tests.

208, **TITLE:** Detection of Covid-19 Through the Analysis of Vocal Fold Oscillations  
<https://doi.org/10.1109/ICASSP39728.2021.9414201>  
**AUTHORS:** M. Al Ismail, S. Deshmukh and R. Singh  
**HIGHLIGHT:** Our goal is to validate this hypothesis, and to quantitatively characterize the changes observed to enable the detection of COVID-19 from voice.

209, **TITLE:** Ct-Caps: Feature Extraction-Based Automated Framework for Covid-19 Disease Identification From Chest Ct Scans Using Capsule Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414214>  
**AUTHORS:** S. Heidarian et al.  
**HIGHLIGHT:** In this paper, a Capsule network framework, referred to as the "CT-CAPS", is presented to automatically extract distinctive features of chest CT scans.

210, **TITLE:** Few-Shot Learning for Ct Scan Based Covid-19 Diagnosis  
<https://doi.org/10.1109/ICASSP39728.2021.9413443>  
**AUTHORS:** Y. Jiang, H. Chen, H. Ko and D. K. Han  
**HIGHLIGHT:** In order to tackle the above issues, we propose a supervised domain adaption based COVID-19 CT diagnostic method which can perform effectively when only a small samples of labeled CT scans are available.

211, **TITLE:** Graph-Based Pyramid Global Context Reasoning With a Saliency- Aware Projection for Covid-19 Lung Infections Segmentation  
<https://doi.org/10.1109/ICASSP39728.2021.9413957>  
**AUTHORS:** H. Huang et al.  
**HIGHLIGHT:** To tackle these issues, we propose a Graph-based Pyramid Global Context Reasoning (Graph-PGCR) module, which is capable of modeling long-range dependencies among disjoint infections as well as adapt size variation.

212, **TITLE:** Interpreting Glottal Flow Dynamics for Detecting Covid-19 From Voice  
<https://doi.org/10.1109/ICASSP39728.2021.9414530>  
**AUTHORS:** S. Deshmukh, M. Al Ismail and R. Singh  
**HIGHLIGHT:** This paper proposes a method that analyzes the differential dynamics of the glottal flow waveform (GFW) during voice production to identify features in them that are most significant for the detection of COVID-19 from voice.

213, **TITLE:** Cycle Generative Adversarial Network Approaches to Produce Novel Portable Chest X-Rays Images for Covid-19 Diagnosis  
<https://doi.org/10.1109/ICASSP39728.2021.9414031>  
**AUTHORS:** D. I. Mor?s, J. de Moura, J. Novo and M. Ortega  
**HIGHLIGHT:** In this work, given the low availability of images of this recent disease, we present new approaches to artificially increase the dimensionality of portable chest X-ray datasets for COVID-19 diagnosis.

214, **TITLE:** EEG-Based Emotion Classification Using Graph Signal Processing  
<https://doi.org/10.1109/ICASSP39728.2021.9414342>  
**AUTHORS:** S. S. Saboksayr, G. Mateos and M. Cetin  
**HIGHLIGHT:** In this work, we bring to bear graph signal processing (GSP) techniques to tackle the problem of automatic emotion recognition using brain signals.

215, **TITLE:** Granger Causality Based Directional Phase-Amplitude Coupling Measure  
<https://doi.org/10.1109/ICASSP39728.2021.9414004>



- AUTHORS: T. T. K. Munia and S. Aviyente  
HIGHLIGHT: In this paper, we introduce a Granger causality (GC) based approach to estimate the direction of PAC.
- 216, TITLE: REPAC: Reliable Estimation of Phase-Amplitude Coupling in Brain Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414749>  
AUTHORS: G. Cisotto  
HIGHLIGHT: This contribution presents REPAC, a reliable and robust algorithm for modeling and detecting PAC events in EEG signals.
- 217, TITLE: Subspace Oddity - Optimization on Product of Stiefel Manifolds for EEG Data  
<https://doi.org/10.1109/ICASSP39728.2021.9413730>  
AUTHORS: M. Sayu Yamamoto, F. Yger and S. Chevallier  
HIGHLIGHT: In this paper, we propose a novel similarity-based classification method that relies on dimensionality reduction of EEG covariance matrices.
- 218, TITLE: Decentralized Motion Inference and Registration of Neuropixel Data  
<https://doi.org/10.1109/ICASSP39728.2021.9414145>  
AUTHORS: E. Varol et al.  
HIGHLIGHT: We introduce a new registration method to partially correct for this motion.
- 219, TITLE: Dynamic Graph Learning Based on Graph Laplacian  
<https://doi.org/10.1109/ICASSP39728.2021.9413744>  
AUTHORS: B. Jiang, Y. Yu, H. Krim and S. L. Smith  
HIGHLIGHT: The purpose of this paper is to infer a global (collective) model of time-varying responses of a set of nodes as a dynamic graph, where the individual time series are respectively observed at each of the nodes.
- 220, TITLE: Mutual Information Flows in a Bivariate Point Process  
<https://doi.org/10.1109/ICASSP39728.2021.9414484>  
AUTHORS: S. Ahmed Pasha and V. Solo  
HIGHLIGHT: Here we address that question using mutual information flows and establish a connection with Granger causality.
- 221, TITLE: Uncertainty-Based Biological Age Estimation of Brain MRI Scans  
<https://doi.org/10.1109/ICASSP39728.2021.9414112>  
AUTHORS: K. Armanious, S. Abdulatif, W. Shi, T. Hepp, S. Gatidis and B. Yang  
HIGHLIGHT: In this initial study, we propose a new framework for organ-specific BA estimation utilizing 3D magnetic resonance image (MRI) scans.
- 222, TITLE: Sparse Representation of Complex-Valued fMRI Data Based on Hard Thresholding of Spatial Source Phase  
<https://doi.org/10.1109/ICASSP39728.2021.9414589>  
AUTHORS: J. -Y. Song, M. -Y. Qi, D. -P. Lv, C. -Y. Zhang, Q. -H. Lin and V. D. Calhoun  
HIGHLIGHT: This study proposes a sparse representation method using SSP hard thresholding to achieve the sparsity of spatial components, enabling the use of initially complex-valued fMRI data and retaining the brain information embedded in noisy voxels and weak BOLD-related voxels with small phase values.
- 223, TITLE: Tucker Decomposition for Extracting Shared and Individual Spatial Maps from Multi-Subject Resting-State fMRI Data  
<https://doi.org/10.1109/ICASSP39728.2021.9413958>  
AUTHORS: Y. Han, Q. -H. Lin, L. -D. Kuang, X. -F. Gong, F. Cong and V. D. Calhoun  
HIGHLIGHT: This study proposes to decompose multi-subject fMRI data in a natural three-way of voxel ? time ? subject via TKD.
- 224, TITLE: Riemannian Geometry-Based Decoding of the Directional Focus of Auditory Attention Using EEG  
<https://doi.org/10.1109/ICASSP39728.2021.9413404>  
AUTHORS: S. Geirnaert, T. Francart and A. Bertrand  
HIGHLIGHT: Here, we propose Riemannian geometry-based classification (RGC) as an alternative for this CSP approach, in which the covariance matrix of a new EEG segment is directly classified while taking its Riemannian structure into account.
- 225, TITLE: DFDM: A Deep Feature Decoupling Module for Lung Nodule Segmentation  
<https://doi.org/10.1109/ICASSP39728.2021.9414938>

AUTHORS: W. Chen, Q. Wang, S. Huang, X. Zhang, Y. Li and C. Liu  
HIGHLIGHT: In this paper, we propose a novel feature decoupling method to tackle two critical problems in the lung nodule segmentation task: (i) ambiguity of nodule boundary leads to the imprecise segmentation boundary and (ii) the high false positive rate of segmentation result.

226, TITLE: Pyramid U-Net for Retinal Vessel Segmentation  
<https://doi.org/10.1109/ICASSP39728.2021.9414164>  
AUTHORS: J. Zhang, Y. Zhang and X. Xu  
HIGHLIGHT: In this paper, we propose pyramid U-Net for accurate retinal vessel segmentation.

227, TITLE: A Probabilistic Model for Segmentation of Ambiguous 3D Lung Nodule  
<https://doi.org/10.1109/ICASSP39728.2021.9415006>  
AUTHORS: X. Long et al.  
HIGHLIGHT: To this end we propose a probabilistic generative segmentation model consisting of a V-Net and a conditional variational autoencoder.

228, TITLE: Semi-Supervised Skin Lesion Segmentation with Learning Model Confidence  
<https://doi.org/10.1109/ICASSP39728.2021.9414297>  
AUTHORS: Z. Xie, E. Tu, H. Zheng, Y. Gu and J. Yang  
HIGHLIGHT: In this paper, to solve this issue, we propose a novel confidence aware semi-supervised learning method based on a mean teacher scheme.

229, TITLE: A Hybrid Feature Enhancement Method for GI And Segmentation In Histopathology Images  
<https://doi.org/10.1109/ICASSP39728.2021.9415095>  
AUTHORS: X. Wu, X. Li, K. Hu, Z. Chen and X. Gao  
HIGHLIGHT: In this paper, a hybrid feature enhancement network (HFE-Net) for glandular segmentation is proposed, which includes a multi-scale local feature extraction block (MSLFEB) and a global feature enhancement block (GFEB).

230, TITLE: Automated Multi-Organ Segmentation in Pet Images Using Cascaded Training of a 3d U-Net and Convolutional Autoencoder  
<https://doi.org/10.1109/ICASSP39728.2021.9414700>  
AUTHORS: A. Liebgott, C. Lorenz, S. Gatidis, V. C. Vu, K. Nikolaou and B. Yang  
HIGHLIGHT: As this transfer of information from CT/MRI to the PET domain is not always feasible, e.g. when the corresponding CT or MRI images are unavailable or corrupted by artifacts, we propose a novel approach to perform organ segmentation on the PET images directly.

231, TITLE: Improved Supervised Training of Physics-Guided Deep Learning Image Reconstruction with Multi-Masking  
<https://doi.org/10.1109/ICASSP39728.2021.9413495>  
AUTHORS: B. Yaman, S. A. H. Hosseini, S. Moeller and M. Ak?akaya  
HIGHLIGHT: In this study, we propose to improve the performance and robustness of supervised training by utilizing randomness by retrospectively selecting only a subset of all the available measurements for data consistency units.

232, TITLE: Fine-Grained Mri Reconstruction Using Attentive Selection Generative Adversarial Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414981>  
AUTHORS: J. Liu and M. Yaghoobi  
HIGHLIGHT: Inspired by the state-of-the-art methods in image generation, we propose a novel attention-based deep learning framework to provide high-quality MRI reconstruction.

233, TITLE: Ensure: Ensemble Stein's Unbiased Risk Estimator for Unsupervised Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9414513>  
AUTHORS: H. K. Aggarwal, A. Pramanik and M. Jacob  
HIGHLIGHT: We propose an ENsemble SURE (ENSURE) approach to train a deep network only from undersampled measurements.

234, TITLE: Ultrasound Elasticity Imaging Using Physics-Based Models and Learning-Based Plug-and-Play Priors  
<https://doi.org/10.1109/ICASSP39728.2021.9413652>  
AUTHORS: N. Mohammadi, M. M. Dooley and M. Cetin  
HIGHLIGHT: Integrating learning-based priors with physical forward models for ultrasound elasticity imaging, we present a joint reconstruction framework which guarantees that learning driven reconstructions are consistent with the underlying physics.

- 235, TITLE: A Periodic Frame Learning Approach for Accurate Landmark Localization in M-Mode Echocardiography  
<https://doi.org/10.1109/ICASSP39728.2021.9414375>  
AUTHORS: Y. Tian, S. Xu, L. Guo and F. Cong  
HIGHLIGHT: In this paper, we propose a novel two-stage frame-level detection and heatmap regression model for accurate landmark localization in m-mode echocardiography, which promotes better integration between global context information and local appearance.
- 236, TITLE: A Bias-Reducing Loss Function for CT Image Denoising  
<https://doi.org/10.1109/ICASSP39728.2021.9413855>  
AUTHORS: M. Nagare, R. Melnyk, O. Rahman, K. D. Sauer and C. A. Bouman  
HIGHLIGHT: In this paper, we present a novel approach to designing a loss function that penalizes variance and bias differently.
- 237, TITLE: Learning Binary Semantic Embedding for Breast Histology Image Classification and Retrieval  
<https://doi.org/10.1109/ICASSP39728.2021.9415036>  
AUTHORS: X. Kang, X. Liu, X. Nie and Y. Yin  
HIGHLIGHT: To address this issues, we propose a novel method for Learning Binary Semantic Embedding (LBSE).
- 238, TITLE: Channel Attention Residual U-Net for Retinal Vessel Segmentation  
<https://doi.org/10.1109/ICASSP39728.2021.9414282>  
AUTHORS: C. Guo, M. Szemenyei, Y. Hu, W. Wang, W. Zhou and Y. Yi  
HIGHLIGHT: In this work, we propose a new deep learning model, namely Channel Attention Residual U-Net (CAR-UNet), to accurately segment retinal vascular and non-vascular pixels.
- 239, TITLE: CMIM: Cross-Modal Information Maximization For Medical Imaging  
<https://doi.org/10.1109/ICASSP39728.2021.9414132>  
AUTHORS: T. Sylvain et al.  
HIGHLIGHT: In this paper, we propose an innovative framework that makes the most of available data by learning good representations of a multi-modal input that are resilient to modality dropping at test-time, using recent advances in mutual information maximization.
- 240, TITLE: Structure-Enhanced Attentive Learning For Spine Segmentation From Ultrasound Volume Projection Images  
<https://doi.org/10.1109/ICASSP39728.2021.9414658>  
AUTHORS: R. Zhao et al.  
HIGHLIGHT: In this paper, we propose a novel framework to improve the segmentation accuracy on spine images via structure-enhanced attentive learning.
- 241, TITLE: Foveal Avascular Zone Segmentation of Octa Images Using Deep Learning Approach with Unsupervised Vessel Segmentation  
<https://doi.org/10.1109/ICASSP39728.2021.9415070>  
AUTHORS: Z. Liang, J. Zhang and C. An  
HIGHLIGHT: To simultaneously implement vessel and accurate FAZ segmentation, an end-to-end trained network is proposed to achieve unsupervised vessel segmentation and supervised FAZ segmentation.
- 242, TITLE: Acute Lymphoblastic Leukemia Detection Based on Adaptive Unsharpening and Deep Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9414362>  
AUTHORS: A. Genovese, M. S. Hosseini, V. Piuri, K. N. Plataniotis and F. Scotti  
HIGHLIGHT: To address this issue, in this paper we propose the first machine learning-based approach able to enhance blood sample images by an adaptive unsharpening method.
- 243, TITLE: Meta Ordinal Weighting Net For Improving Lung Nodule Classification  
<https://doi.org/10.1109/ICASSP39728.2021.9413622>  
AUTHORS: Y. Lei, H. Shan and J. Zhang  
HIGHLIGHT: In this paper, we propose a Meta Ordinal Weighting Network (MOW-Net) to explicitly align each training sample with a meta ordinal set (MOS) containing a few samples from all classes.
- 244, TITLE: Deepnodule: Multi-Task Learning of Segmentation Bootstrap for Pulmonary Nodule Detection  
<https://doi.org/10.1109/ICASSP39728.2021.9413825>  
AUTHORS: J. Li, K. Wang, D. Yang, X. Zhang and C. Liu  
HIGHLIGHT: To overcome those barriers, we present a novel multi-task 3D convolutional network (DeepNodule) for simultaneous nodule detection and segmentation in a shared-and-fined manner.

- 245, TITLE: Dense Attention Module for Accurate Pulmonary Nodule Detection  
<https://doi.org/10.1109/ICASSP39728.2021.9413936>  
AUTHORS: J. Liu, J. Li, F. Xue and C. Wu  
HIGHLIGHT: In this paper, we propose a novel pulmonary nodule detection framework and a novel 3D dense attention module (DAM) which can efficiently exploit the abundant 3D spatial features.
- 246, TITLE: Unsupervised Multimodal Image Registration with Adaptive Gradient Guidance  
<https://doi.org/10.1109/ICASSP39728.2021.9414320>  
AUTHORS: Z. Xu, J. Yan, J. Luo, X. Li and J. Jagadeesan  
HIGHLIGHT: In this paper, we propose a novel multimodal registration framework, which leverages the deformation fields estimated from both: (i) the original to-be-registered image pair, (ii) their corresponding gradient intensity maps, and adaptively fuses them with the proposed gated fusion module.
- 247, TITLE: Improving Intraoperative Liver Registration in Image-Guided Surgery with Learning-Based Reconstruction  
<https://doi.org/10.1109/ICASSP39728.2021.9414549>  
AUTHORS: M. Jia and M. Kyan  
HIGHLIGHT: To overcome the problems caused by noisy, partial, and sparse intraoperative sampling, we propose a novel occupancy-learning-based mesh to point cloud registration and apply it to align the preoperative liver image to intraoperative samples.
- 248, TITLE: A New Framework Based on Transfer Learning for Cross-Database Pneumonia Detection  
<https://doi.org/10.1109/ICASSP39728.2021.9414997>  
AUTHORS: X. Shan and Y. Wen  
HIGHLIGHT: In this paper, we proposed a new framework based on transfer learning for cross-database pneumonia detection.
- 249, TITLE: Hierarchical Attention-Based Temporal Convolutional Networks for Eeg-Based Emotion Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9413635>  
AUTHORS: C. Li, B. Chen, Z. Zhao, N. Cummins and B. W. Schuller  
HIGHLIGHT: In order to tackle these problems, a hierarchical attention-based temporal convolutional networks (HATCN) for efficient EEG-based emotion recognition is proposed.
- 250, TITLE: Deep Multiway Canonical Correlation Analysis For Multi-Subject Eeg Normalization  
<https://doi.org/10.1109/ICASSP39728.2021.9414274>  
AUTHORS: J. R. Katthi and S. Ganapathy  
HIGHLIGHT: In this paper, we propose a deep learning framework to improve the correlation of electroencephalography (EEG) data recorded from multiple subjects engaged in an audio listening task.
- 251, TITLE: Dynamic Graph Modeling Of Simultaneous EEG And Eye-Tracking Data For Reading Task Identification  
<https://doi.org/10.1109/ICASSP39728.2021.9414343>  
AUTHORS: P. Mathur, T. Mittal and D. Manocha  
HIGHLIGHT: We present a new approach, that we call AdaGTCN, for identifying human reader intent from Electroencephalogram (EEG) and Eye movement (EM) data in order to help differentiate between normal reading and task-oriented reading.
- 252, TITLE: Learning From Heterogeneous Eeg Signals with Differentiable Channel Reordering  
<https://doi.org/10.1109/ICASSP39728.2021.9413712>  
AUTHORS: A. Saeed, D. Grangier, O. Pietquin and N. Zeghidour  
HIGHLIGHT: We propose CHARM, a method for training a single neural network across inconsistent input channels.
- 253, TITLE: Enhancing Multi-Channel Eeg Classification with Gramian Temporal Generative Adversarial Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414078>  
AUTHORS: C. N. Enoch Kan, R. J. Povinelli and D. H. Ye  
HIGHLIGHT: In this paper, we propose a novel method to synthesize multi-channel EEG in the form of Gramian Angular Field (GAF) images with a Gramian Temporal Generative Adversarial Network (GT-GAN).
- 254, TITLE: A Novel Convolutional Neural Network Model to Remove Muscle Artifacts from EEG  
<https://doi.org/10.1109/ICASSP39728.2021.9414228>  
AUTHORS: H. Zhang, C. Wei, M. Zhao, Q. Liu and H. Wu  
HIGHLIGHT: Here we introduce a novel convolutional neural network (CNN) with gradually ascending feature dimensions and downsampling in time series for removing muscle artifacts in EEG data.

- 255, TITLE: Multilabel 12-Lead Electrocardiogram Classification Using Beat to Sequence Autoencoders  
<https://doi.org/10.1109/ICASSP39728.2021.9414934>  
 AUTHORS: A. W. Wong, A. Salimi, A. Hindle, S. V. Kalmady and P. Kaul  
 HIGHLIGHT: This paper investigates the multi-label, multi-class classification of ECG records into one or more of 27 possible medical diagnoses.
- 256, TITLE: Contrastive Embedding Learning Method for Respiratory Sound Classification  
<https://doi.org/10.1109/ICASSP39728.2021.9414385>  
 AUTHORS: W. Song, J. Han and H. Song  
 HIGHLIGHT: To address the problems, we propose a contrastive embedding learning method, where the input is a contrastive tuple.
- 257, TITLE: Decoding Neural Representations of Rhythmic Sounds From Magnetoencephalography  
<https://doi.org/10.1109/ICASSP39728.2021.9414178>  
 AUTHORS: P. -C. Chang et al.  
 HIGHLIGHT: In this work, we investigate how to extract rhythmic information embedded in the brain responses and to decode the original audio waveforms from the extracted information.
- 258, TITLE: Low-Dimensional Denoising Embedding Transformer for ECG Classification  
<https://doi.org/10.1109/ICASSP39728.2021.9413766>  
 AUTHORS: J. Guan, W. Wang, P. Feng, X. Wang and W. Wang  
 HIGHLIGHT: In this paper, we propose a new method for ECG classification, called low-dimensional denoising embedding transformer (LDTF), which contains two components, i.e., low-dimensional denoising embedding (LDE) and transformer learning.
- 259, TITLE: Self-Supervised Learning for Sleep Stage Classification with Predictive and Discriminative Contrastive Coding  
<https://doi.org/10.1109/ICASSP39728.2021.9414752>  
 AUTHORS: Q. Xiao et al.  
 HIGHLIGHT: The purpose of this paper is to learn efficient representations from raw electroencephalogram (EEG) signals for sleep stage classification via self-supervised learning (SSL).
- 260, TITLE: Length No Longer Matters: A Real Length Adaptive Arrhythmia Classification Model with Multi-Scale Convolution  
<https://doi.org/10.1109/ICASSP39728.2021.9414616>  
 AUTHORS: C. Han, F. Yu, P. Wang, R. Huang, X. Huang and L. Cui  
 HIGHLIGHT: To address these problems, we propose a length adaptive arrhythmia classification model that can take advantage of raw ECG records of variable length.
- 261, TITLE: Few-Shot Learning for Decoding Surface Electromyography for Hand Gesture Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9413582>  
 AUTHORS: E. Rahimian, S. Zabihi, A. Asif, S. F. Atashzar and A. Mohammadi  
 HIGHLIGHT: Therefore, in this work, we develop a novel hand gesture recognition framework based on the formulation of FewShot Learning (FSL) to infer the required output given only one or a few numbers of training examples.
- 262, TITLE: Deep Lung Auscultation Using Acoustic Biomarkers for Abnormal Respiratory Sound Event Detection  
<https://doi.org/10.1109/ICASSP39728.2021.9414845>  
 AUTHORS: U. Tiwari, S. Bhosale, R. Chakraborty and S. K. Kopparapu  
 HIGHLIGHT: In this paper, we propose to use two sets of diversified acoustic biomarkers extracted using Discrete Wavelet Transform (DWT) and deep encoded features from the intermediate layer of a pre-trained Audio Event Detection (AED) model trained using sounds from daily activities.
- 263, TITLE: Speaker-Independent Brain Enhanced Speech Denoising  
<https://doi.org/10.1109/ICASSP39728.2021.9414969>  
 AUTHORS: M. Hosseini, L. Celotti and ? Plourde  
 HIGHLIGHT: In this paper, we propose a novel deep learning method referred to as the Brain Enhanced Speech Denoiser (BESD), that takes advantage of the attended auditory information present in the brain activity of the listener to denoise a multi-talker speech.
- 264, TITLE: Shapelet Based Visual Assessment of Cluster Tendency in Analyzing Complex Upper Limb Motion  
<https://doi.org/10.1109/ICASSP39728.2021.9414144>

- AUTHORS: S. Datta, C. Karmakar, P. Rathore and M. Palaniswami  
HIGHLIGHT: We propose an unsupervised method for shapelet extraction using maximin shape sampling and shape-based distance computation for selecting key shapelets representing characteristic motion patterns.
- 265, TITLE: Human-Centered Favorite Music Classification Using EEG-Based Individual Music Preference Via Deep Time-Series CCA  
<https://doi.org/10.1109/ICASSP39728.2021.9414945>  
AUTHORS: R. Sawata, T. Ogawa and M. Haseyama  
HIGHLIGHT: A method to classify a user's like or dislike musical pieces based on the extraction of his or her music preference is proposed in this paper.
- 266, TITLE: Multi-Scale and Multi-Region Facial Discriminative Representation for Automatic Depression Level Prediction  
<https://doi.org/10.1109/ICASSP39728.2021.9413504>  
AUTHORS: M. Niu, J. Tao and B. Liu  
HIGHLIGHT: For these reasons, we propose a multi-scale and multi-region facial dynamic representation method to improve the prediction performance.
- 267, TITLE: ECG Heart-Beat Classification Using Multimodal Image Fusion  
<https://doi.org/10.1109/ICASSP39728.2021.9414709>  
AUTHORS: Z. Ahmad, A. Tabassum, L. Guan and N. Khan  
HIGHLIGHT: In this paper, we present a novel Image Fusion Model (IFM) for ECG heart-beat classification to overcome the weaknesses of existing machine learning techniques that rely either on manual feature extraction or direct utilization of 1D raw ECG signal.
- 268, TITLE: Estimation of Visual Features of Viewed Image From Individual and Shared Brain Information Based on FMRI Data Using Probabilistic Generative Model  
<https://doi.org/10.1109/ICASSP39728.2021.9414507>  
AUTHORS: T. Higashi, K. Maeda, T. Ogawa and M. Haseyama  
HIGHLIGHT: This paper presents a method for estimation of visual features based on brain responses measured when subjects view images.
- 269, TITLE: Hierarchical Pose Classification for Infant Action Analysis and Mental Development Assessment  
<https://doi.org/10.1109/ICASSP39728.2021.9415088>  
AUTHORS: J. Zhou, Z. Jiang, J. -H. Yoo and J. -N. Hwang  
HIGHLIGHT: This paper presents a hierarchical pose classifier, given a baby image frame that combines the benefits of 3D human pose estimation and scene context information.
- 270, TITLE: On The Relationship Between Speech-Based Breathing Signal Prediction Evaluation Measures and Breathing Parameters Estimation  
<https://doi.org/10.1109/ICASSP39728.2021.9414756>  
AUTHORS: Z. Mostaani, V. Srikanth Nallanthighal, A. H?rm?, H. Strik and M. Magimai-Doss  
HIGHLIGHT: This paper investigates whether there is a systematic relationship between the different objective measures used for training and evaluating the neural network models and the end-goal, i.e. estimation of breathing parameters such as, breathing rate and tidal volume.
- 271, TITLE: Prediction of Egfr Mutation Status in Lung Adenocarcinoma Using Multi-Source Feature Representations  
<https://doi.org/10.1109/ICASSP39728.2021.9414064>  
AUTHORS: J. Cheng, J. Liu, M. Jiang, H. Yue, L. Wu and J. Wang  
HIGHLIGHT: In this study, we propose a hybrid framework, namely HC-DLR, to noninvasively predict EGFR mutation status by fusing multi-source features including low-level handcrafted radiomics (HCR) features, high-level deep learning-based radiomics (DLR) features, and demographics features.
- 272, TITLE: Training Neural Networks with Domain Pattern-Aware Auxiliary Task for Sleep Staging  
<https://doi.org/10.1109/ICASSP39728.2021.9413683>  
AUTHORS: T. Lee, J. Hwang and H. Lee  
HIGHLIGHT: Accordingly, we present an auxiliary classification task for sleep staging to enable NNs to exploit clinically significant EEG patterns in data.
- 273, TITLE: Classification of Expert-Novice Level Using Eye Tracking And Motion Data via Conditional Multimodal Variational Autoencoder  
<https://doi.org/10.1109/ICASSP39728.2021.9414361>

- AUTHORS: Y. Akamatsu, K. Maeda, T. Ogawa and M. Haseyama  
HIGHLIGHT: In this paper, we propose a semi-supervised anomaly detection approach that requires only sensor data of experts for training and identifies those of novices as anomalies.
- 274, TITLE: Gate Trimming: One-Shot Channel Pruning for Efficient Convolutional Neural Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414338>  
AUTHORS: F. Yu, C. Han, P. Wang, X. Huang and L. Cui  
HIGHLIGHT: In this work, we propose a one-shot global pruning approach called Gate Trimming (GT), which is more efficient to compress the CNNs.
- 275, TITLE: Deep S3PR: Simultaneous Source Separation and Phase Retrieval Using Deep Generative Models  
<https://doi.org/10.1109/ICASSP39728.2021.9413714>  
AUTHORS: C. A. Metzler and G. Wetzstein  
HIGHLIGHT: In this work, we demonstrate that by restricting the solutions to lie in the range of a deep generative model, we can constrain the search space sufficiently to solve S3PR. Code associated with this work is available at <https://github.com/computational-imaging/DeepS3PR>.
- 276, TITLE: Adversarial Attacks on Object Detectors with Limited Perturbations  
<https://doi.org/10.1109/ICASSP39728.2021.9414125>  
AUTHORS: Z. Shi et al.  
HIGHLIGHT: In this paper, we present a novel attack framework named DTTACK to fool both one-stage and two-stage object detectors with limited perturbations.
- 277, TITLE: A Consensus Equilibrium Solution For Deep Image Prior Powered By Red  
<https://doi.org/10.1109/ICASSP39728.2021.9414290>  
AUTHORS: R. Hyder, H. Mansour, Y. Ma, P. T. Boufounos and P. Wang  
HIGHLIGHT: In this paper, we formulate DeepRED as a consensus equilibrium problem and set up a fixed-point algorithm for solving the equilibrium equations.
- 278, TITLE: Suremap: Predicting Uncertainty in Cnn-Based Image Reconstructions Using Stein's Unbiased Risk Estimate  
<https://doi.org/10.1109/ICASSP39728.2021.9414306>  
AUTHORS: R. Kitichotkul, C. A. Metzler, F. Ong and G. Wetzstein  
HIGHLIGHT: In this work we use Stein's unbiased risk estimate (SURE) to develop per-pixel confidence intervals, in the form of heatmaps, for compressive sensing reconstruction using the approximate message passing (AMP) framework with CNN-based denoisers.
- 279, TITLE: Multi-Initialization Meta-Learning with Domain Adaptation  
<https://doi.org/10.1109/ICASSP39728.2021.9414554>  
AUTHORS: Z. Chen and D. Wang  
HIGHLIGHT: To improve the performance on multi-modal tasks, we propose multi-initialization meta-learning with domain adaptation (MIML-DA) to tackle such domain shift.
- 280, TITLE: Stochastic Deep Unfolding for Imaging Inverse Problems  
<https://doi.org/10.1109/ICASSP39728.2021.9414332>  
AUTHORS: J. Liu, Y. Sun, W. Gan, X. Xu, B. Wohlberg and U. S. Kamilov  
HIGHLIGHT: We propose SCRED-Net as a novel methodology that introduces a stochastic approximation to the unfolded regularization by denoising (RED) algorithm.
- 281, TITLE: Fusion-Based Digital Image Correlation Framework for Strain Measurement  
<https://doi.org/10.1109/ICASSP39728.2021.9414987>  
AUTHORS: L. Shi, D. Liu, M. Umeda and N. Hana  
HIGHLIGHT: To overcome this issue, we propose an end-to-end DIC framework incorporating the image fusion principle to achieve full-field strain measurement over the curved surface.
- 282, TITLE: Learning Sparsifying Transforms for Image Reconstruction in Electrical Impedance Tomography  
<https://doi.org/10.1109/ICASSP39728.2021.9413861>  
AUTHORS: K. Yang, N. Borjindargoon, B. P. Ng, S. Ravishankar and B. Wen  
HIGHLIGHT: We propose a blind compressed sensing algorithm, dubbed TL-EIT, which simultaneously optimizes the sparsifying transform and updates the reconstructed image.



- 283, TITLE: D-VDAMP: Denoising-Based Approximate Message Passing for Compressive MRI  
<https://doi.org/10.1109/ICASSP39728.2021.9414708>  
AUTHORS: C. A. Metzler and G. Wetzstein  
HIGHLIGHT: In this work, we propose a CNN architecture for removing colored Gaussian noise and combine it with the recently proposed VDAMP algorithm, whose effective noise follows a predictable colored Gaussian distribution.
- 284, TITLE: Empirically Accelerating Scaled Gradient Projection Using Deep Neural Network for Inverse Problems in Image Processing  
<https://doi.org/10.1109/ICASSP39728.2021.9415023>  
AUTHORS: B. H. Lee and S. Y. Chun  
HIGHLIGHT: Here, we present a novel DNN-based convergent iterative algorithm that accelerates conventional optimization algorithms.
- 285, TITLE: Synthetic Aperture Acoustic Imaging with Deep Generative Model Based Source Distribution Prior  
<https://doi.org/10.1109/ICASSP39728.2021.9414138>  
AUTHORS: B. Fan and S. Das  
HIGHLIGHT: In this work, we propose to image large acoustic sources with a combination of synthetic aperture and their geometric structures modeled by a conditional generative adversarial network (cGAN).
- 286, TITLE: Non-Local Single Image DE-Raining Without Decomposition  
<https://doi.org/10.1109/ICASSP39728.2021.9415000>  
AUTHORS: C. Zheng, Z. Li, Y. Li and S. Wu  
HIGHLIGHT: On top of a new insight in single image de-raining, a nonlocal de-raining algorithm is proposed in this paper to remove the rain streaks from the rainy image.
- 287, TITLE: Frame-Rate-Aware Aggregation for Efficient Video Super-Resolution  
<https://doi.org/10.1109/ICASSP39728.2021.9414334>  
AUTHORS: T. Isobe, F. Zhu and S. Wang  
HIGHLIGHT: In contrast to the previous works that perform explicit motion estimation and compensation, we propose a novel deep neural network which performs implicit motion estimation with frame-rate-based temporal aggregation.
- 288, TITLE: Measurement Coding Framework with Adjacent Pixels Based Measurement Matrix for Compressively Sensed Images  
<https://doi.org/10.1109/ICASSP39728.2021.9414068>  
AUTHORS: R. Wan, J. Zhou, B. Huang, H. Zeng and Y. Fan  
HIGHLIGHT: To further compress measurements, the output of block-based compressed sensing, this work presents a measurement coding framework using measurement-domain intra prediction.
- 289, TITLE: Multiview Sensing with Unknown Permutations: an Optimal Transport Approach  
<https://doi.org/10.1109/ICASSP39728.2021.9415075>  
AUTHORS: Y. Ma, P. T. Boufounos, H. Mansour and S. Aeron  
HIGHLIGHT: In this paper we take a fresh look at this problem through the lens of optimal transport (OT).
- 290, TITLE: A High-Frame-Rate Eye-Tracking Framework for Mobile Devices  
<https://doi.org/10.1109/ICASSP39728.2021.9414624>  
AUTHORS: Y. Chang, C. He, Y. Zhao, T. Lu and N. Gu  
HIGHLIGHT: In this work, we tackle the tracking efficiency challenge and introduce GazeHFR, a biologic-inspired eye-tracking model specialized for mobile devices, offering both high accuracy and efficiency.
- 291, TITLE: Catiloc: Camera Image Transformer for Indoor Localization  
<https://doi.org/10.1109/ICASSP39728.2021.9414939>  
AUTHORS: A. Ghofrani, R. M. Toroghi and S. Mojtaba Tabatabaie  
HIGHLIGHT: In this paper the problem of single image indoor camera localization has been addressed.
- 292, TITLE: Sar Image Autofocusing Using Wirtinger Calculus and Cauchy Regularization  
<https://doi.org/10.1109/ICASSP39728.2021.9413629>  
AUTHORS: Z. -Y. Zhang, O. Pappas and A. Achim  
HIGHLIGHT: In this paper, an optimization model using Cauchy regularization is proposed for simultaneous SAR image reconstruction and autofocusing.

- 293, TITLE: A Homogeneity-Based Multiscale Hyperspectral Image Representation for Sparse Spectral Unmixing  
<https://doi.org/10.1109/ICASSP39728.2021.9415024>  
AUTHORS: L. C. Ayres, S. J. M. de Almeida, J. C. M. Bermudez and R. A. Borsoi  
HIGHLIGHT: In this work, we propose a computationally efficient multiscale representation method for hyperspectral data adapted to the unmixing problem.
- 294, TITLE: Learning to Estimate Kernel Scale and Orientation of Defocus Blur with Asymmetric Coded Aperture  
<https://doi.org/10.1109/ICASSP39728.2021.9413920>  
AUTHORS: J. Li, Q. Dai and J. Wen  
HIGHLIGHT: To tackle this problem, we propose a deep-learning-based framework estimating the kernel scale and orientation of the defocus blur to adjust lens focus rapidly.
- 295, TITLE: Transmittance Regularizer for Binary coded Aperture Design in a Computational Imaging end-to-end Approach  
<https://doi.org/10.1109/ICASSP39728.2021.9413903>  
AUTHORS: J. Bacca, T. Gelvez and H. Arguello  
HIGHLIGHT: Therefore, this work proposes two transmittance regularizers that jointly induce binary en-tries and adjust the transmittance level to be incorporated in an E2E approach.
- 296, TITLE: Fourier Transformation Autoencoders for Anomaly Detection  
<https://doi.org/10.1109/ICASSP39728.2021.9415010>  
AUTHORS: D. Lappas, V. Argyriou and D. Makris  
HIGHLIGHT: This paper introduces Fourier Trans-forms into AutoEncoders to demonstrate how the inclusion of a frequency domain presents less noisy features for a deep learning network to detect anomalies.
- 297, TITLE: Zero-Gradient Constraints for Destriping of Remote-Sensing Data  
<https://doi.org/10.1109/ICASSP39728.2021.9415066>  
AUTHORS: K. Naganuma, S. Takeyama and S. Ono  
HIGHLIGHT: This paper proposes an effective and efficient destriping method for remote-sensing data.
- 298, TITLE: Selection Based on Statistical Characteristics for Object Detection  
<https://doi.org/10.1109/ICASSP39728.2021.9413848>  
AUTHORS: Z. Li, Y. Yuan and D. Ma  
HIGHLIGHT: In this paper, we propose a multi-scale sample selection based on statistical characteristics for object detection.
- 299, TITLE: CSPN: Multi-Scale Cascade Spatial Pyramid Network for Object Detection  
<https://doi.org/10.1109/ICASSP39728.2021.9414883>  
AUTHORS: T. Wang, C. Ma, H. Su and W. Wang  
HIGHLIGHT: To solve this problem and obtain better detection performance, we propose a novel net-work named Multi-Scale Cascade Spatial Pyramid Network (MS-CSPN) to strengthen Feature Pyramid.
- 300, TITLE: Dual-Stream Network Based On Global Guidance for Salient Object Detection  
<https://doi.org/10.1109/ICASSP39728.2021.9413702>  
AUTHORS: S. Gao, Q. Guo, W. Zhang, W. Zhang and Z. Ji  
HIGHLIGHT: To remedy the problems, we propose a dual-stream network based on global guidance with two plug-ins, global attention based multi-scale high-level feature extraction module (GAMS) to mine global guidance and scale adaptive global guidance module (SAGG) to seamlessly integrate the global guidance into each decoding layer.
- 301, TITLE: SSFENet: Spatial and Semantic Feature Enhancement Network for Object Detection  
<https://doi.org/10.1109/ICASSP39728.2021.9413602>  
AUTHORS: T. Wang, C. Ma, H. Su and W. Wang  
HIGHLIGHT: In this paper, we present a novel network to address this dilemma, denoted as Spatial and Semantic Feature Enhancement Network (SSFENet).
- 302, TITLE: Saliency-Driven Versatile Video Coding for Neural Object Detection  
<https://doi.org/10.1109/ICASSP39728.2021.9415048>  
AUTHORS: K. Fischer, F. Fleckenstein, C. Herglotz and A. Kaup  
HIGHLIGHT: In this paper, we pro-pose such a saliency-driven coding framework for the video coding for machines task using the latest video coding standard Versatile Video Coding (VVC).
- 303, TITLE: Object-Oriented Relational Distillation for Object Detection

<https://doi.org/10.1109/ICASSP39728.2021.9413925>

AUTHORS: S. Miao and R. Feng

HIGHLIGHT: To this end, we propose a novel Object-Oriented Relational Distillation (OORD) method that drives small detection models to have an effective performance like large detection models with constant efficiency.

304, TITLE: Ensembling Object Detectors for Image and Video Data Analysis

<https://doi.org/10.1109/ICASSP39728.2021.9414013>

AUTHORS: K. Chumachenko, J. Raitoharju, A. Iosifidis and M. Gabbouj

HIGHLIGHT: In this paper, we propose a method for ensembling the outputs of multiple object detectors for improving detection performance and precision of bounding boxes on image data.

305, TITLE: Training Real-Time Panoramic Object Detectors with Virtual Dataset

<https://doi.org/10.1109/ICASSP39728.2021.9414503>

AUTHORS: Q. -Y. Shen, T. -G. Huang, P. -X. Ding and J. He

HIGHLIGHT: In this paper, we propose a panoramic virtual dataset for training object detectors on 360° images.

306, TITLE: Fast: Feature Aggregation for Detecting Salient Object in Real-Time

<https://doi.org/10.1109/ICASSP39728.2021.9414457>

AUTHORS: L. Tang, B. Li, Y. Wu, B. Xiao and S. Ding

HIGHLIGHT: This paper introduces a method named FAST for real-time salient object detection with an extremely efficient CNN architecture.

307, TITLE: Exploiting the Dual-Tree Complex Wavelet Transform for Ship Wake Detection in SAR Imagery

<https://doi.org/10.1109/ICASSP39728.2021.9414898>

AUTHORS: W. Ma, A. Achim and O. Karakus

HIGHLIGHT: In this paper, we analyse synthetic aperture radar (SAR) images of the sea surface using an inverse problem formulation whereby Radon domain information is enhanced in order to accurately detect ship wakes.

308, TITLE: Task-Related Self-Supervised Learning For Remote Sensing Image Change Detection

<https://doi.org/10.1109/ICASSP39728.2021.9414387>

AUTHORS: Z. Cai, Z. Jiang and Y. Yuan

HIGHLIGHT: In this work, an unsupervised change detection method based on Task-related Self-supervised Learning Change Detection network with smooth mechanism(TSLCD) is proposed to eliminate it.

309, TITLE: Unsupervised Common Particular Object Discovery and Localization by Analyzing a Match Graph

<https://doi.org/10.1109/ICASSP39728.2021.9413549>

AUTHORS: M. Okuda, S. Satoh, Y. Sato and Y. Kidawara

HIGHLIGHT: This paper describes an unsupervised method that more accurately discovers and localizes common particular objects within a set of images.

310, TITLE: Predictive Coding for Lossless Dataset Compression

<https://doi.org/10.1109/ICASSP39728.2021.9413447>

AUTHORS: M. Barowsky, A. Mariona and F. P. Calmon

HIGHLIGHT: We prove the equivalence of dataset compression to compressing a permutation-invariant structure of the data and implement such a scheme via predictive coding.

311, TITLE: Adaptive Dual Tree Structure For Screen Content Coding

<https://doi.org/10.1109/ICASSP39728.2021.9414002>

AUTHORS: W. Zhu, J. Xu, L. Zhang and Y. Wang

HIGHLIGHT: Therefore, adaptive dual tree structure is proposed in this paper wherein the coding structure of each coding tree unit is switched between separate and joint coding structure to adapt the textures adaptively.

312, TITLE: SNR-Adaptive Deep Joint Source-Channel Coding for Wireless Image Transmission

<https://doi.org/10.1109/ICASSP39728.2021.9414037>

AUTHORS: M. Ding, J. Li, M. Ma and X. Fan

HIGHLIGHT: Considering the problem of joint source-channel coding (JSCC) for multi-user transmission of images over noisy channels, an autoencoder-based novel deep joint source-channel coding scheme is proposed in this paper.

313, TITLE: Relying on a Rate Constraint to Reduce Motion Estimation Complexity

<https://doi.org/10.1109/ICASSP39728.2021.9414799>

AUTHORS: G. B. Sant'Anna, L. Henrique Cancellier, I. Seidel, M. Grellert and J. L. G?ntzel  
HIGHLIGHT: This paper proposes a rate-based candidate elimination strategy for Motion Estimation, which is considered one of the main sources of encoder complexity.

314, TITLE: A Novel Viewport-Adaptive Motion Compensation Technique for Fisheye Video  
<https://doi.org/10.1109/ICASSP39728.2021.9413576>  
AUTHORS: A. Regensky, C. Herglotz and A. Kaup  
HIGHLIGHT: We propose a novel viewport-adaptive motion compensation technique that applies the motion vectors in different perspective viewports in order to realize these motion planes.

315, TITLE: Rate-Distortion Optimized Motion Estimation for on-the-Sphere Compression of 360 Videos  
<https://doi.org/10.1109/ICASSP39728.2021.9413681>  
AUTHORS: A. Marie, N. Mahmoudian Bidgoli, T. Maugey and A. Roumy  
HIGHLIGHT: In this paper, the on-the-sphere compression [1] for omnidirectional still images is extended to videos.

316, TITLE: Adaptive GOP Size Decision for Multi-Pass Video Coding Based on Hidden Markov Model  
<https://doi.org/10.1109/ICASSP39728.2021.9414161>  
AUTHORS: B. Li, J. Han and Y. Xu  
HIGHLIGHT: In this paper, a novel method to determine the size of each group of picture (GOP) using the multi-pass information is presented.

317, TITLE: Improved Intra Mode Coding Beyond Av1  
<https://doi.org/10.1109/ICASSP39728.2021.9413420>  
AUTHORS: Y. Jin, L. Zhao, X. Zhao, S. Liu and A. C. Bovik  
HIGHLIGHT: In this paper, two methods are proposed to further reduce the signaling cost of delta angles: cross-component delta angle coding, and context-adaptive delta angle coding, whereby the cross-component and spatial correlation of the delta angles are explored, respectively.

318, TITLE: Decision Tree Based Inter Partition Termination For Av1 Encoding  
<https://doi.org/10.1109/ICASSP39728.2021.9413481>  
AUTHORS: X. Chen, Y. Zhang, Y. Li and J. Wen  
HIGHLIGHT: To deal with this problem, in this paper, we propose a decision tree based algorithm to early terminate the inter prediction process by predicting splitting decisions at each depth.

319, TITLE: Image Coding For Machines: an End-To-End Learned Approach  
<https://doi.org/10.1109/ICASSP39728.2021.9414465>  
AUTHORS: N. Le, H. Zhang, F. Cricri, R. Ghaznavi-Youvalari and E. Rahtu  
HIGHLIGHT: In this paper, we propose an image codec for machines which is neural network (NN) based and end-to-end learned.

320, TITLE: Sparse Flow Adversarial Model For Robust Image Compression  
<https://doi.org/10.1109/ICASSP39728.2021.9415100>  
AUTHORS: S. Zhao, S. Yang, Z. Liu, Z. Feng and X. Liu  
HIGHLIGHT: In this paper, we propose a novel compression method called sparse flow adversarial model (SFAM).

321, TITLE: HVS-Based Perceptual Color Compression of Image Data  
<https://doi.org/10.1109/ICASSP39728.2021.9414773>  
AUTHORS: L. Prangnell and V. Sanchez  
HIGHLIGHT: In this paper, we propose a novel perceptual image coding technique, named Perceptual Color Compression (PCC).

322, TITLE: HOCA: Higher-Order Channel Attention for Single Image Super-Resolution  
<https://doi.org/10.1109/ICASSP39728.2021.9414892>  
AUTHORS: Y. Lv, T. Dai, B. Chen, J. Lu, S. -T. Xia and J. Cao  
HIGHLIGHT: To address this issue, we propose a higher-order channel attention (HOCA) module to enhance the representation ability of CNNs.

323, TITLE: Image Super-Resolution Using Multi-Resolution Attention Network  
<https://doi.org/10.1109/ICASSP39728.2021.9414535>  
AUTHORS: A. Liu, S. Li and Y. Chang

**HIGHLIGHT:** To address these issues, we propose a multi-resolution attention network (MRAN), which progressively reconstructs images at large scale factors by aggregating features from multiple resolutions.

324, **TITLE:** Real Image Super-Resolution Using Token Based Contextual Attention

<https://doi.org/10.1109/ICASSP39728.2021.9414593>

**AUTHORS:** Z. Pan and B. Li

**HIGHLIGHT:** To alleviate these issues, we propose a new token based attention module with innovative contextual encoding to enable SR models to be robust to image patch sizes at testing.

325, **TITLE:** Feature Redundancy Mining: Deep Light-Weight Image Super-Resolution Model

<https://doi.org/10.1109/ICASSP39728.2021.9413846>

**AUTHORS:** J. Xiao, W. Jia and K. -M. Lam

**HIGHLIGHT:** In this paper, by considering the correlation and redundancy of feature maps, we propose a feature information mining network to efficiently investigate the features, for the SISR problem.

326, **TITLE:** Lightweight Non-Local Network for Image Super-Resolution

<https://doi.org/10.1109/ICASSP39728.2021.9414527>

**AUTHORS:** R. Wang, T. Lei, W. Zhou, Q. Wang, H. Meng and A. K. Nandi

**HIGHLIGHT:** To address these issues, we propose a lightweight non-local network (LNLN) for image super resolution in this paper.

327, **TITLE:** Lightweight and Accurate Single Image Super-Resolution with Channel Segregation Network

<https://doi.org/10.1109/ICASSP39728.2021.9414039>

**AUTHORS:** Z. -H. Niu, X. -P. Lin, A. -N. Yu, Y. -H. Zhou and Y. -B. Yang

**HIGHLIGHT:** To address this issue, we propose an efficient channel segregation block containing multiple branches with different depths, enabling the model to preserve basic content, and focusing on optimizing the detail content with fewer parameters.

328, **TITLE:** Deep Learning Architectural Designs for Super-Resolution Of Noisy Images

<https://doi.org/10.1109/ICASSP39728.2021.9414733>

**AUTHORS:** A. Villar-Corrales, F. Schirmacher and C. Riess

**HIGHLIGHT:** In this work, we propose to jointly perform denoising and super-resolution.

329, **TITLE:** Joint Coupled Transform Learning Framework for Multimodal Image Super-Resolution

<https://doi.org/10.1109/ICASSP39728.2021.9413490>

**AUTHORS:** A. Gigie, A. A. Kumar, A. Majumdar, K. Kumar and M. G. Chandra

**HIGHLIGHT:** In this paper, we model the cross-modal dependencies between different modalities for Multimodal Image Super-Resolution (MISR), i.e., enhance the Low Resolution (LR) image of target modality with the guidance of a High Resolution (HR) image from another modality.

330, **TITLE:** Hyperspectral Image Super-Resolution Via Adjacent Spectral Fusion Strategy

<https://doi.org/10.1109/ICASSP39728.2021.9413980>

**AUTHORS:** Q. Li, Q. Wang and X. Li

**HIGHLIGHT:** To address this issue, we explore a new structure for hyperspectral image SR via adjacent spectral fusion strategy.

331, **TITLE:** Raw Data Processing for Practical Time-of-Flight Super-Resolution

<https://doi.org/10.1109/ICASSP39728.2021.9414354>

**AUTHORS:** M. H. Conde

**HIGHLIGHT:** In this paper we note that while attempting to address the last two issues, e. g., via burst mode, the lateral resolution can be effectively increased.

332, **TITLE:** Edge-Aware Multi-Scale Progressive Colorization

<https://doi.org/10.1109/ICASSP39728.2021.9414764>

**AUTHORS:** J. Xia, G. Tan, Y. Xiao, F. Xu and C. -S. Leung

**HIGHLIGHT:** To address these problems, we propose a novel edge-aware multi-scale progressive network (EMSPN).

333, **TITLE:** Learning Representation of Multi-Scale Object for Fine-Grained Image Retrieval

<https://doi.org/10.1109/ICASSP39728.2021.9414308>

**AUTHORS:** K. Sun and J. Zhu

**HIGHLIGHT:** In our work, to extract more local features, we propose a method that could propose multiple discriminative regions on different scales, which could provide more refined local and multi-scale representation for fine-grained image retrieval.

334, **TITLE:** Super-Resolution and Infection Edge Detection Co-Guided Learning for Covid-19 Ct Segmentation  
<https://doi.org/10.1109/ICASSP39728.2021.9414327>

**AUTHORS:** Y. Sang, J. Sun, S. Wang, H. Qi and K. Li

**HIGHLIGHT:** In this paper, we propose a novel super-resolution and infection edge detection co-guided learning network for COVID-19 CT segmentation (CogSeg).

335, **TITLE:** Gating Feature Dense Network for Single Anisotropic Mr Image Super-Resolution  
<https://doi.org/10.1109/ICASSP39728.2021.9414646>

**AUTHORS:** W. He, Y. Hu, L. Wang, Z. He and J. Du

**HIGHLIGHT:** In this work, we propose a gating feature dense network to reconstruct HR MR images from low resolution acquisitions, where we use local residual dense block (LRDB) as the backbone.

336, **TITLE:** Adaptable Ensemble Distillation  
<https://doi.org/10.1109/ICASSP39728.2021.9415015>

**AUTHORS:** Y. Wang, D. Yang, W. Zhang, Z. Jiang and W. Zhang

**HIGHLIGHT:** In this paper, we propose Adaptable Ensemble Distillation (AED) that inherits the merits of existing OKD methods while overcoming their major drawbacks.

337, **TITLE:** A Scale Invariant Measure of Flatness for Deep Network Minima  
<https://doi.org/10.1109/ICASSP39728.2021.9413771>

**AUTHORS:** A. Rangamani, N. H. Nguyen, A. Kumar, D. Phan, S. P. Chin and T. D. Tran

**HIGHLIGHT:** In this paper we show that for deep networks with positively homogenous activations, these rescalings constitute equivalence relations, and that these equivalence relations induce a quotient manifold structure in the parameter space.

338, **TITLE:** Multi-Order Adversarial Representation Learning for Composed Query Image Retrieval  
<https://doi.org/10.1109/ICASSP39728.2021.9414436>

**AUTHORS:** Z. Fu, X. Chen, J. Dong and S. Ji

**HIGHLIGHT:** So this paper proposes a new Multi-order Adversarial Network (MAN) which uses multilevel representations and simultaneously explores their low-order and high-order interactions, obtaining low-order and high-order features.

339, **TITLE:** Deep Neural Networks with Flexible Complexity While Training Based on Neural Ordinary Differential Equations  
<https://doi.org/10.1109/ICASSP39728.2021.9413916>

**AUTHORS:** Z. Luo, S. -i. Kamata, Z. Sun and W. Zhou

**HIGHLIGHT:** In this work, we experimentally investigate the effectiveness of using neural ordinary differential equations (NODEs) as a component to provide further depth to relatively shallower networks rather than stacked layers (depth) which achieved improvement with fewer parameters.

340, **TITLE:** Improving Memory Banks for Unsupervised Learning with Large Mini-Batch, Consistency and Hard Negative Mining  
<https://doi.org/10.1109/ICASSP39728.2021.9414389>

**AUTHORS:** A. Bulat, E. S?nchez-Lozano and G. Tzimiropoulos

**HIGHLIGHT:** In this paper, we introduce 3 improvements to the vanilla memory bank-based formulation which brings massive accuracy gains: (a) Large mini-batch: we pull multiple augmentations for each sample within the same batch and show that this leads to better models and enhanced memory bank updates.

341, **TITLE:** Robust Binary Loss for Multi-Category Classification with Label Noise  
<https://doi.org/10.1109/ICASSP39728.2021.9414493>

**AUTHORS:** D. Liu, G. Yang, J. Wu, J. Zhao and F. Lv

**HIGHLIGHT:** To address the problem, we propose to train deep models with robust binary loss functions.

342, **TITLE:** A Plug and Play Fast Intersection Over Union Loss for Boundary Box Regression  
<https://doi.org/10.1109/ICASSP39728.2021.9413641>

**AUTHORS:** Z. Kuang, X. Fang, R. Zhang, X. Shao and H. Wang

**HIGHLIGHT:** In this paper, we design a Fast Intersection over Union (FIoU) loss, which can not only keep the advantages but also solve the weakness of IoU-based losses.

- 343, TITLE: Attribute Decomposition for Flow-Based Domain Mapping  
<https://doi.org/10.1109/ICASSP39728.2021.9414028>  
AUTHORS: S. -J. Huang and J. -T. Chien  
HIGHLIGHT: To handle the mixed features for better generation, this paper presents an attribute decomposition based on the sequence data and carries out the flow-based image domain mapping.
- 344, TITLE: Ada-Sise: Adaptive Semantic Input Sampling for Efficient Explanation of Convolutional Neural Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414942>  
AUTHORS: M. Sudhakar, S. Sattarzadeh, K. N. Plataniotis, J. Jang, Y. Jeong and H. Kim  
HIGHLIGHT: In this work, we combine both approaches as a hybrid visual explanation algorithm and propose an efficient interpretation method for convolutional neural networks.
- 345, TITLE: Network Pruning Using Linear Dependency Analysis on Feature Maps  
<https://doi.org/10.1109/ICASSP39728.2021.9414393>  
AUTHORS: H. Pan, Z. Chao, J. Qian, B. Zhuang, S. Wang and J. Xiao  
HIGHLIGHT: In this paper, we regard a channel 'redundant' if its output is linearly dependent with respect to those of other channels.
- 346, TITLE: Multiple-Input Multiple-Output Fusion Network for Generalized Zero-Shot Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9413509>  
AUTHORS: F. Zhong, G. Wang, Z. Chen, X. Yuan and F. Xia  
HIGHLIGHT: To address this issue, we propose a Multiple-Input Multiple-Output Fusion Network to GZSL.
- 347, TITLE: Representative Local Feature Mining for Few-Shot Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9413890>  
AUTHORS: K. Yan, L. Liu, J. Hou and P. Wang  
HIGHLIGHT: Given this, we propose a novel method that chooses representative local features to facilitate few-shot learning.
- 348, TITLE: KAN: Knowledge-Augmented Networks for Few-Shot Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9413612>  
AUTHORS: Z. Zhu and X. Lin  
HIGHLIGHT: Therefore, considering that semantic information can enhance understanding when visual information is limited, we propose Knowledge-Augmented Networks (KAN), which combines the visual features with the semantic information extracted from knowledge graph to represent the features of each class.
- 349, TITLE: Few-Shot Image Classification with Multi-Facet Prototypes  
<https://doi.org/10.1109/ICASSP39728.2021.9414374>  
AUTHORS: K. Yan, Z. Bouraoui, P. Wang, S. Jameel and S. Schockaert  
HIGHLIGHT: In particular, we propose an adaptive similarity measure, relying on predicted facet importance weights for a given set of categories.
- 350, TITLE: Self-Supervised Learning for Few-Shot Image Classification  
<https://doi.org/10.1109/ICASSP39728.2021.9413783>  
AUTHORS: D. Chen, Y. Chen, Y. Li, F. Mao, Y. He and H. Xue  
HIGHLIGHT: In this paper, we proposed to train a more generalized embedding network with self-supervised learning (SSL) which can provide robust representation for downstream tasks by learning from the data itself.
- 351, TITLE: Domain Adaptation for Learning Generator From Paired Few-Shot Data  
<https://doi.org/10.1109/ICASSP39728.2021.9414702>  
AUTHORS: C. -C. Teng, P. -Y. Chen and W. -C. Chiu  
HIGHLIGHT: We propose a Paired Few-shot GAN (PFS-GAN) model for learning generators with sufficient source data and a few target data.
- 352, TITLE: Deep Semi-Supervised Metric Learning Via Identification of Manifold Memberships  
<https://doi.org/10.1109/ICASSP39728.2021.9414447>  
AUTHORS: F. Zhuang and P. Moulin  
HIGHLIGHT: We propose a method which allows the use of class-representative anchors (proxies), and avoids the computational costs associated with triplet sampling.
- 353, TITLE: A Ranked Similarity Loss Function with pair Weighting for Deep Metric Learning



<https://doi.org/10.1109/ICASSP39728.2021.9414668>

AUTHORS: J. Wang, Z. Zhang, D. Huang, W. Song, Q. Wei and X. Li

HIGHLIGHT: To address these problems, we propose to build a ranked similarity loss function with pair weighting (dubbed RMS loss).

354, TITLE: Statistical Distance Metric Learning for Image Set Retrieval

<https://doi.org/10.1109/ICASSP39728.2021.9413393>

AUTHORS: T. -Y. Hu and A. G. Hauptmann

HIGHLIGHT: In this paper, we obviate the need of feature aggregation and propose a novel Statistical Distance Metric Learning (SDML) framework, which represents each image set as a probability distribution in embedding feature space and compares two image sets by statistical distance between their distributions.

355, TITLE: Distribution-Aware Hierarchical Weighting Method for Deep Metric Learning

<https://doi.org/10.1109/ICASSP39728.2021.9414864>

AUTHORS: Y. Zhu, Y. Feng, M. Zhou, B. Qiang, L. Hou U and J. Zhu

HIGHLIGHT: In this paper, we propose distribution-aware hierarchical weighting (DHW) method for deep metric learning.

356, TITLE: Integrated Grad-Cam: Sensitivity-Aware Visual Explanation of Deep Convolutional Networks Via Integrated Gradient-Based Scoring

<https://doi.org/10.1109/ICASSP39728.2021.9415064>

AUTHORS: S. Sattarzadeh, M. Sudhakar, K. N. Plataniotis, J. Jang, Y. Jeong and H. Kim

HIGHLIGHT: Addressing this problem, we introduce a solution to tackle this issue by computing the path integral of the gradient-based terms in Grad-CAM.

357, TITLE: Visualizing Association in Exemplar-Based Classification

<https://doi.org/10.1109/ICASSP39728.2021.9413574>

AUTHORS: T. Kashima, R. Hataya and H. Nakayama

HIGHLIGHT: To solve this problem, we propose a novel method of explainable classification; this method uses images representing each image class, which we call exemplars.

358, TITLE: HFGCNET: High-Frequency Graph Reasoning for Finer Semantic Image Segmentation

<https://doi.org/10.1109/ICASSP39728.2021.9413469>

AUTHORS: Z. Sun, R. Wang, Z. Luo and W. Chen

HIGHLIGHT: This work presents a high-frequency graph convolution operation to solve the above problems.

359, TITLE: Unsupervised Image Segmentation with Spatial Triplet Markov Trees

<https://doi.org/10.1109/ICASSP39728.2021.9414435>

AUTHORS: H. Gangloff, J. -B. Courbot, E. Monfrini and C. Collet

HIGHLIGHT: In this article we study an extension of HMTs called Spatial Triplet Markov Trees (STMTs) which is designed to greatly increase the correlations of the random variables while keeping the possibility of direct and exact inference procedures.

360, TITLE: Cross Scene Video Foreground Segmentation Via Co-Occurrence Probability Oriented Supervised and Unsupervised Model Interaction

<https://doi.org/10.1109/ICASSP39728.2021.9413965>

AUTHORS: D. Liang, B. Kang, X. Liu, H. Sun, L. Zhang and N. Liu

HIGHLIGHT: In this paper, we propose a cross scene video foreground segmentation framework to extend the generalization capability of those supervised model depending on scene-specific training.

361, TITLE: Instance Segmentation with the Number of Clusters Incorporated in Embedding Learning

<https://doi.org/10.1109/ICASSP39728.2021.9414312>

AUTHORS: J. Cao and H. Yan

HIGHLIGHT: In this work, we propose to embed prior clustering information into an embedding learning framework FCRNet, stimulating the one-stage instance segmentation.

362, TITLE: Decouple the High-Frequency and Low-Frequency Information of Images for Semantic Segmentation

<https://doi.org/10.1109/ICASSP39728.2021.9414019>

AUTHORS: L. Shan, X. Li and W. Wang

HIGHLIGHT: At present, the semantic segmentation methods are all based on CNN and ignore the advantages of traditional image processing technology. We combine the two and make them promote each other.

- 363, TITLE: MPDNet: A 3D Missing Part Detection Network Based on Point Cloud Segmentation  
<https://doi.org/10.1109/ICASSP39728.2021.9414867>  
AUTHORS: Z. Fan, H. Liu, J. He, M. Zhang and X. Du  
HIGHLIGHT: To tackle the issue, in this paper, we propose a novel model named MPDNet, which exploits 3D point cloud pairs as input for missing part detection.
- 364, TITLE: SM+: Refined Scale Match for Tiny Person Detection  
<https://doi.org/10.1109/ICASSP39728.2021.9414162>  
AUTHORS: N. Jiang, X. Yu, X. Peng, Y. Gong and Z. Han  
HIGHLIGHT: In this paper, we investigate the scale alignment between pre-training and target datasets, and propose a new refined Scale Match method (termed SM+) for tiny person detection.
- 365, TITLE: Sub-Band Grouping Spectral Feature-Attention Block for Hyperspectral Image Classification  
<https://doi.org/10.1109/ICASSP39728.2021.9414678>  
AUTHORS: W. Zhou, S. -i. Kamata and Z. Luo  
HIGHLIGHT: In this paper, we proposed a novel sub-band grouping recurrent neural network (RNN) model with gated recurrent units (GRUs) to find the intrinsic feature in spectral information.
- 366, TITLE: Unsupervised Stacked Capsule Autoencoder for Hyperspectral Image Classification  
<https://doi.org/10.1109/ICASSP39728.2021.9413664>  
AUTHORS: E. Pan, Y. Ma, X. Mei, F. Fan and J. Ma  
HIGHLIGHT: Based on HSI data's aforementioned structural characteristics, combined with the stacked capsule autoencoder, we propose our model to achieve an unsupervised HSI classification.
- 367, TITLE: Robust Graph Autoencoder for Hyperspectral Anomaly Detection  
<https://doi.org/10.1109/ICASSP39728.2021.9414767>  
AUTHORS: G. Fan, Y. Ma, J. Huang, X. Mei and J. Ma  
HIGHLIGHT: In order to tackle these problems, we propose a robust graph autoencoder (RGAE) for hyperspectral anomaly detection.
- 368, TITLE: Reflectance-Oriented Probabilistic Equalization for Image Enhancement  
<https://doi.org/10.1109/ICASSP39728.2021.9414651>  
AUTHORS: X. Wu, Y. Sun, A. Kimura and K. Kashino  
HIGHLIGHT: To solve this problem, we propose a novel 2D histogram equalization approach.
- 369, TITLE: PD-GAN: Perceptual-Details GAN for Extremely Noisy Low Light Image Enhancement  
<https://doi.org/10.1109/ICASSP39728.2021.9413433>  
AUTHORS: Y. Liu, Z. Wang, Y. Zeng, H. Zeng and D. Zhao  
HIGHLIGHT: To solve the problems, we propose perceptual-details GAN (PD-GAN) utilizing Zero-DCE to initially recover illumination and combine residual dense-block Encoder-Decoder structure to suppress noise while finely adjusting the illumination.
- 370, TITLE: Heterogeneous two-Stream Network with Hierarchical Feature Prefusion for Multispectral Pan-Sharpener  
<https://doi.org/10.1109/ICASSP39728.2021.9413736>  
AUTHORS: D. Wang, Y. Bai, B. Bai, C. Wu and Y. Li  
HIGHLIGHT: In this paper, we propose a heterogeneous two-stream network (HTSNet) with hierarchical feature prefusion for MS pan-sharpening.
- 371, TITLE: Synergic Feature Attention for Image Restoration  
<https://doi.org/10.1109/ICASSP39728.2021.9413484>  
AUTHORS: C. Mou and J. Zhang  
HIGHLIGHT: To overcome these problems, in this paper, we propose a novel Synergic Attention Network (SAT-Net) for image restoration as an inventive attempt to combine local and non-local attention mechanisms to restore complex textures and highly repetitive details distinguishingly.
- 372, TITLE: Efficient Multi-Objective GANs for Image Restoration  
<https://doi.org/10.1109/ICASSP39728.2021.9413897>  
AUTHORS: J. Su and H. Yin  
HIGHLIGHT: Here, we propose an efficient formulation of multiple loss components for training GANs.
- 373, TITLE: Self-Convolution: A Highly-Efficient Operator for Non-Local Image Restoration

<https://doi.org/10.1109/ICASSP39728.2021.9414124>

AUTHORS: L. Guo, Z. Zha, S. Ravishankar and B. Wen

HIGHLIGHT: In this work, we propose a novel Self-Convolution operator to exploit image non-local similarity in a self-supervised way.

374, TITLE: NMF-SAE: An Interpretable Sparse Autoencoder for Hyperspectral Unmixing

<https://doi.org/10.1109/ICASSP39728.2021.9414084>

AUTHORS: F. Xiong, J. Zhou, M. Ye, J. Lu and Y. Qian

HIGHLIGHT: In this paper, we combine the advantages of both model-based and learning-based methods and propose a nonnegative matrix factorization (NMF) inspired sparse autoencoder (NMF-SAE) for hyperspectral unmixing.

375, TITLE: An ADMM Based Network for Hyperspectral Unmixing Tasks

<https://doi.org/10.1109/ICASSP39728.2021.9414555>

AUTHORS: C. Zhou and M. R. D. Rodrigues

HIGHLIGHT: In this paper, we use algorithm unrolling approaches in order to design a new neural network structure applicable to hyperspectral unmixing challenges.

376, TITLE: Variational Autoencoders for Hyperspectral Unmixing with Endmember Variability

<https://doi.org/10.1109/ICASSP39728.2021.9414940>

AUTHORS: S. Shi, M. Zhao, L. Zhang and J. Chen

HIGHLIGHT: This paper presents a variational autoencoder (VAE) framework for hyperspectral unmixing accounting for the endmember variability.

377, TITLE: Augmented Gaussian Linear Mixture Model for Spectral Variability in Hyperspectral Unmixing

<https://doi.org/10.1109/ICASSP39728.2021.9414358>

AUTHORS: Y. E. Salehani, E. Arabnejad and S. Gazor

HIGHLIGHT: In this paper, we propose a novel hyperspectral unmixing through the perturbed linear mixture model to take into account the spectral variability offset of the linear mixture model.

378, TITLE: UTDN: An Unsupervised Two-Stream Dirichlet-Net for Hyperspectral Unmixing

<https://doi.org/10.1109/ICASSP39728.2021.9414810>

AUTHORS: Q. Jin, Y. Ma, X. Mei, H. Li and J. Ma

HIGHLIGHT: In this paper, we propose a novel two-stream Dirichlet-net, termed as uTDN, to address the above problems.

379, TITLE: Laplacian Regularized Tensor Low-Rank Minimization for Hyperspectral Snapshot Compressive Imaging

<https://doi.org/10.1109/ICASSP39728.2021.9413381>

AUTHORS: Y. Yang, F. Jiang and H. Lu

HIGHLIGHT: In this paper, we propose a tensor-based low-rank reconstruction algorithm with hyper-Laplacian constraint for hyperspectral SCI systems.

380, TITLE: Compressing Local Descriptor Models for Mobile Applications

<https://doi.org/10.1109/ICASSP39728.2021.9414416>

AUTHORS: R. Miles and K. Mikołajczyk

HIGHLIGHT: In this paper, we consider these practical aspects and improve the state-of-the-art HardNet model through the use of depthwise separable layers and an efficient tensor decomposition.

381, TITLE: VK-Net: Category-Level Point Cloud Registration with Unsupervised Rotation Invariant Keypoints

<https://doi.org/10.1109/ICASSP39728.2021.9414384>

AUTHORS: Z. Chen, W. Yang, Z. Xu, Z. Shi and L. Huang

HIGHLIGHT: In this paper, we propose VK-Net, a neural network that learns to discover a set of category-specific keypoints from a single point cloud in an unsupervised manner.

382, TITLE: Matching as Color Images: Thermal Image Local Feature Detection and Description

<https://doi.org/10.1109/ICASSP39728.2021.9414341>

AUTHORS: B. Deshpande, S. Hanamsheth, Y. Lu and G. Lu

HIGHLIGHT: Motivated by this challenge, we propose a triplet based Siamese CNN for feature detection and extraction for any given thermal image.

383, TITLE: Frame Rate Up-Conversion Using Key Point Agnostic Frequency-Selective Mesh-to-Grid Resampling

<https://doi.org/10.1109/ICASSP39728.2021.9413684>

AUTHORS: V. Heimann, A. Spruck and A. Kaup  
HIGHLIGHT: We use the model-based key point agnostic frequency-selective mesh-to-grid resampling (AFSMR) for this task and show that AFSMR works best for applications that contain irregular meshes with varying densities.

384, TITLE: Efficient Real-Time Video Stabilization with a Novel Least Squares Formulation  
<https://doi.org/10.1109/ICASSP39728.2021.9414545>  
AUTHORS: J. Ke, A. J. Watras, J. -J. Kim, H. Liu, H. Jiang and Y. H. Hu  
HIGHLIGHT: We present a novel video stabilization algorithm (LSstab) that removes unwanted motions in real-time.

385, TITLE: Decomposing Textures using Exponential Analysis  
<https://doi.org/10.1109/ICASSP39728.2021.9413909>  
AUTHORS: Y. Hou, A. Cuyt, W. -s. Lee and D. Bhowmik  
HIGHLIGHT: We present a new approach using a recent 2-dimensional exponential analysis technique.

386, TITLE: G-Arrays: Geometric Arrays for Efficient Point Cloud Processing  
<https://doi.org/10.1109/ICASSP39728.2021.9413902>  
AUTHORS: H. Roodaki, M. Dehyadegari and M. N. Bojnordi  
HIGHLIGHT: In this paper, we propose a novel data structure for representing point clouds with a reduced memory requirement and a faster lookup than the state-of-the-art formats.

387, TITLE: QoE-Driven and Tile-Based Adaptive Streaming for Point Clouds  
<https://doi.org/10.1109/ICASSP39728.2021.9414121>  
AUTHORS: L. Wang, C. Li, W. Dai, J. Zou and H. Xiong  
HIGHLIGHT: To address this, we propose a QoE-driven and tile-based adaptive streaming approach for point clouds, to reduce transmission redundancy and maximize user's QoE.

388, TITLE: Dynamic Point Cloud Compression Using A Cuboid Oriented Discrete Cosine Based Motion Model  
<https://doi.org/10.1109/ICASSP39728.2021.9414171>  
AUTHORS: A. Ahmmed, M. Paul, M. Murshed and D. Taubman  
HIGHLIGHT: An improved commonality modeling technique is proposed that employs discrete cosine basis oriented motion models and the domains of such models are approximated by homogeneous regions called cuboids.

389, TITLE: An Adaptive Pyramid Single-View Depth Lookup Table Coding Method  
<https://doi.org/10.1109/ICASSP39728.2021.9414584>  
AUTHORS: Y. Cai, R. Wang, S. Gu, J. Zhang and W. Gao  
HIGHLIGHT: In this paper, an adaptive pyramid single-view depth lookup table coding method is proposed, with the purpose of designing a clean syntax structure in the sequence header with reasonably good performance.

390, TITLE: Patch Decoder-Side Depth Estimation In Mpeg Immersive Video  
<https://doi.org/10.1109/ICASSP39728.2021.9414056>  
AUTHORS: M. Milovanovic, F. Henry, M. Cagnazzo and J. Jung  
HIGHLIGHT: This paper presents a new approach for achieving bitrate and pixel rate reduction in the MPEG immersive video coding setting.

391, TITLE: Geometry Consistency Of Augmented Reality Based On Semantics  
<https://doi.org/10.1109/ICASSP39728.2021.9414088>  
AUTHORS: H. Quan, M. Yao and X. Qian  
HIGHLIGHT: In augmented reality, for achieving geometric consistency in the perspective projection virtual-real, we propose a semantic consistency method to achieve the fusion between virtual and real objects with selected segmented objects in the real scene as references.

392, TITLE: What And Where To Focus In Person Search  
<https://doi.org/10.1109/ICASSP39728.2021.9414439>  
AUTHORS: T. Zhou and K. Tian  
HIGHLIGHT: For these two findings, we first propose multilevel semantic aggregation algorithm for more discriminative feature descriptors. Then, a pose-assisted attention module is designed to highlight fine-grained area of the target and simultaneously capture valuable clues for identification.

393, TITLE: Stable and Effective One-Step Method for Person Search  
<https://doi.org/10.1109/ICASSP39728.2021.9413460>

AUTHORS: N. Lv, X. Xiang, X. Wang, J. Yang, R. Abdeen and A. El Saddik  
HIGHLIGHT: In this work, we propose an end-to-end model containing the feature extractor, the region proposal network, and the multi-task learning module.

394, TITLE: An Adaptive Part-Based Model For Person Re-Identification  
<https://doi.org/10.1109/ICASSP39728.2021.9415086>  
AUTHORS: X. -P. Lin and Y. -B. Yang  
HIGHLIGHT: To address the part-misalignment problem and learn a more discriminative embedding for person Re-ID, we propose a novel Adaptive Part-based Model (APM), which adaptively partition the extracted feature maps by a Partition-Aware module to learn an embedding.

395, TITLE: Crowd Counting Via Multi-Level Regression With Latent Gaussian Maps  
<https://doi.org/10.1109/ICASSP39728.2021.9414256>  
AUTHORS: Y. Gao and H. Yang  
HIGHLIGHT: In this paper, a novel end-to-end crowd counting framework via multi-level regression with latent Gaussian maps is proposed, which is consisted of GaussianNet, EstimateNet and Discriminator.

396, TITLE: Lightweight Dual-Task Networks For Crowd Counting In Aerial Images  
<https://doi.org/10.1109/ICASSP39728.2021.9413949>  
AUTHORS: Y. Tian, C. Duan, R. Zhang, Z. Wei and H. Wang  
HIGHLIGHT: Therefore, this paper proposes a lightweight dual-task network (LDNet) for crowd counting, which only uses bifurcated structure to overcome these new challenges in aerial images without complicated pipelines.

397, TITLE: SANet++: Enhanced Scale Aggregation with Densely Connected Feature Fusion for Crowd Counting  
<https://doi.org/10.1109/ICASSP39728.2021.9415091>  
AUTHORS: S. Pan, Y. Zhao, F. Su and Z. Zhao  
HIGHLIGHT: In this paper, we present SANet++ with a novel architecture to generate high-quality density maps and further perform accurate counting.

398, TITLE: Attentive Semantic Exploring for Manipulated Face Detection  
<https://doi.org/10.1109/ICASSP39728.2021.9414225>  
AUTHORS: Z. Chen and H. Yang  
HIGHLIGHT: Therefore, we propose a novel manipulated face detection method based on Multilevel Facial Semantic Segmentation and Cascade Attention Mechanism.

399, TITLE: Efficient Face Manipulation Via Deep Feature Disentanglement And Reintegration Net  
<https://doi.org/10.1109/ICASSP39728.2021.9415043>  
AUTHORS: B. Cheng, T. Dai, B. Chen, S. Xia and X. Li  
HIGHLIGHT: To address this issue, we propose a novel Feature Disentanglement and Reintegration network (FDRNet), which employs ground-truth images as informative supervision and dynamically adapts the fusion of informative features of the ground-truth images effectively and efficiently.

400, TITLE: Continuous Face Aging Generative Adversarial Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414429>  
AUTHORS: S. Jeon, P. Lee, K. Hong and H. Byun  
HIGHLIGHT: To this end, we propose the continuous face aging generative adversarial networks (CFA-GAN).

401, TITLE: Fast Inverse Mapping of Face GANs  
<https://doi.org/10.1109/ICASSP39728.2021.9413532>  
AUTHORS: N. Bayat, V. R. Khazaie and Y. Mohsenzadeh  
HIGHLIGHT: We train a ResNet architecture to map given faces to latent vectors that can be used to generate faces nearly identical to the target.

402, TITLE: Multi-Level Adaptive Region of Interest and Graph Learning for Facial Action Unit Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9413551>  
AUTHORS: J. Yan, B. Jiang, J. Wang, Q. Li, C. Wang and S. Pu  
HIGHLIGHT: In this paper, we propose a novel multi-level adaptive ROI and graph learning (MARGL) framework to tackle this problem.

403, TITLE: Bridging Unpaired Facial Photos and Sketches by Line-Drawings

<https://doi.org/10.1109/ICASSP39728.2021.9414380>

AUTHORS: M. Shang, F. Gao, X. Li, J. Zhu and L. Dai

HIGHLIGHT: In this paper, we propose a novel method to learn face sketch synthesis models by using unpaired data.

404, TITLE: Temporal Rain Decomposition with Spatial Structure Guidance for Video Deraining

<https://doi.org/10.1109/ICASSP39728.2021.9413563>

AUTHORS: X. Xue, Y. Ding, L. Ma, Y. Wang, R. Liu and X. Fan

HIGHLIGHT: In this work, we propose a multi-frame deraining network with temporal rain decomposition and spatial structure guidance to more effectively accomplish video deraining.

405, TITLE: GTA-Net: Gradual Temporal Aggregation Network for Fast Video Deraining

<https://doi.org/10.1109/ICASSP39728.2021.9413698>

AUTHORS: X. Xue, X. Meng, L. Ma, R. Liu and X. Fan

HIGHLIGHT: In this work, to effectively exploit temporal information, we develop a simple but effective network, Gradual Temporal Aggregation Network (GTA-Net for short).

406, TITLE: Dense Feature Pyramid Grids Network for Single Image Deraining

<https://doi.org/10.1109/ICASSP39728.2021.9415034>

AUTHORS: Z. Wang, C. Wang, Z. Su and J. Chen

HIGHLIGHT: In this paper, we propose a novel densely connected network with Dense Feature Pyramid Grids Modules, called DFPGN, to solve the rain removal task.

407, TITLE: A Fast and Efficient Network for Single Image Deraining

<https://doi.org/10.1109/ICASSP39728.2021.9414978>

AUTHORS: Y. Yang and H. Lu

HIGHLIGHT: To tackle this problem, we propose a novel Adaptive Dilated Network (ADN) to remove rain streaks from a single image while using less parameters and running faster than previous methods.

408, TITLE: DNANet: Dense Nested Attention Network for Single Image Dehazing

<https://doi.org/10.1109/ICASSP39728.2021.9414179>

AUTHORS: D. Ren, J. Li, M. Han and M. Shu

HIGHLIGHT: In this paper, we propose an innovative approach, called Dense Nested Attention Network (DNANet), to directly restore a clear image from a hazy image with a new topology of connection paths.

409, TITLE: FWB-Net: Front White Balance Network for Color Shift Correction in Single Image Dehazing Via Atmospheric Light Estimation

<https://doi.org/10.1109/ICASSP39728.2021.9414200>

AUTHORS: C. Wang, Y. Huang, Y. Zou and Y. Xu

HIGHLIGHT: Bearing this in mind, in this study, first, a new non-homogeneous atmospheric scattering model (NH-ASM) is proposed for improving image modeling of hazy images taken under complex illumination conditions. Second, a new U-Net based front white balance module (FWB-Module) is dedicatedly designed to correct color shift before generating dehazing result via atmospheric light estimation.

410, TITLE: Learning Integrodifferential Models for Image Denoising

<https://doi.org/10.1109/ICASSP39728.2021.9414914>

AUTHORS: T. Alt and J. Weickert

HIGHLIGHT: We introduce an integrodifferential extension of the edge-enhancing anisotropic diffusion model for image denoising.

411, TITLE: Unrolling of Deep Graph Total Variation for Image Denoising

<https://doi.org/10.1109/ICASSP39728.2021.9414453>

AUTHORS: H. Vu, G. Cheung and Y. C. Eldar

HIGHLIGHT: In this paper, we combine classical graph signal filtering with deep feature learning into a competitive hybrid design?one that utilizes interpretable analytical low-pass graph filters and employs 80% fewer network parameters than state-of-the-art DL denoising scheme DnCNN.

412, TITLE: Learning Model-Blind Temporal Denoisers without Ground Truths

<https://doi.org/10.1109/ICASSP39728.2021.9413606>

AUTHORS: Y. Li, B. Guo, J. Wen, Z. Xia, S. Liu and Y. Han

HIGHLIGHT: In this paper, we propose a general framework for temporal denoising that successfully addresses these challenges.

- 413, TITLE: Image Denoising Based on Correlation Adaptive Sparse Modeling  
<https://doi.org/10.1109/ICASSP39728.2021.9414213>  
AUTHORS: H. Liu, J. Zhang and C. Mou  
HIGHLIGHT: This paper aims to fully exploit local and non-local correlation of image contents separately so that near-optimal sparse representations are achieved and thus the uncertainty of signals is minimized.
- 414, TITLE: NASA: A Noise-Adaptive and Structure-Aware Learning Framework for Image Deblurring  
<https://doi.org/10.1109/ICASSP39728.2021.9414459>  
AUTHORS: X. Liu, L. Ma, R. Liu, W. Zhong, X. Fan and Z. Luo  
HIGHLIGHT: To settle this issue, we develop a Noise-Adaptive Structure-Aware learning framework (NASA) to achieve fully intelligent manufacturing.
- 415, TITLE: Multiple Auxiliary Networks for Single Blind Image Deblurring  
<https://doi.org/10.1109/ICASSP39728.2021.9413674>  
AUTHORS: C. Li, Q. Wang, S. Liu and X. Li  
HIGHLIGHT: In this paper, we propose Multiple Auxiliary Networks (MANet) for single blind image deblurring to assist norm L1-loss function and enhance the quality of the deblurring image.
- 416, TITLE: Joint Learning of Image Aesthetic Quality Assessment and Semantic Recognition Based on Feature Enhancement  
<https://doi.org/10.1109/ICASSP39728.2021.9414367>  
AUTHORS: X. Liu, X. Nie, Z. Shen and Y. Yin  
HIGHLIGHT: In this paper, we explore the relationships between aesthetic quality assessment and semantic recognition task, and employ a multi-task convolutional neural network with feature enhancement mechanism to effectively integrate these two tasks.
- 417, TITLE: Nested Error Map Generation Network for No-Reference Image Quality Assessment  
<https://doi.org/10.1109/ICASSP39728.2021.9413489>  
AUTHORS: J. Chen, H. Wang, G. Li and S. Liu  
HIGHLIGHT: We propose a multi-task learning neural network for No-Reference image quality assessment (NR-IQA).
- 418, TITLE: Regression or classification? New methods to evaluate no-reference picture and video quality models  
<https://doi.org/10.1109/ICASSP39728.2021.9414232>  
AUTHORS: Z. Tu et al.  
HIGHLIGHT: To make the problem more tractable, we propose two new methods - binary, and ordinal classification - as alternatives to evaluate and compare no-reference quality models at coarser levels.
- 419, TITLE: Blind Image Quality Evaluator with Scale Robustness  
<https://doi.org/10.1109/ICASSP39728.2021.9414441>  
AUTHORS: C. Wang and M. Li  
HIGHLIGHT: In this paper, we propose a more generalized blind image quality evaluator with scale robustness (BIQESR) to assess image quality by locating the robust feature points in a multi-scale space.
- 420, TITLE: Multi-Scale Feature-Guided Stereoscopic Video Quality Assessment Based on 3d Convolutional Neural Network  
<https://doi.org/10.1109/ICASSP39728.2021.9414231>  
AUTHORS: Y. Feng, S. Li and Y. Chang  
HIGHLIGHT: In this paper, we proposed a multi-scale feature-guided 3D convolutional neural network for SVQA which not only use 3D convolution to capture spatio-temporal features but also aggregate multi-scale information by a new multi-scale unit.
- 421, TITLE: No-Reference Stereoscopic Image Quality Assessment Based on the Human Visual System  
<https://doi.org/10.1109/ICASSP39728.2021.9414691>  
AUTHORS: F. Meng, S. Li and Y. Chang  
HIGHLIGHT: Recently witnessed the significant progress of biotechnology and motivated by the deeper research on the HVS, we take a step to bridge the gap between HVS and SIQA by generalizing the optic chiasm algorithm and introducing biological vision fusion mechanism in our work.
- 422, TITLE: Stereo Rectification Based on Epipolar Constrained Neural Network  
<https://doi.org/10.1109/ICASSP39728.2021.9413735>  
AUTHORS: Y. Wang, Y. Lu and G. Lu



**HIGHLIGHT:** This paper proposes a novel deep neural network-based method for stereo image rectification.

423, **TITLE:** Multi-Scale Cascade Disparity Refinement Stereo Network

<https://doi.org/10.1109/ICASSP39728.2021.9414923>

**AUTHORS:** X. Jia et al.

**HIGHLIGHT:** Therefore, this paper presents MCDRNet, which combines traditional methods with neural networks to achieve real-time and accurate stereo matching results.

424, **TITLE:** Hierarchical Context Guided Aggregation Network for Stereo Matching

<https://doi.org/10.1109/ICASSP39728.2021.9414381>

**AUTHORS:** J. Peng, W. Xie, Z. Huang, W. Chen and Y. Zhao

**HIGHLIGHT:** In this paper, we propose a simple yet efficient network named Hierarchical Context Guided Aggregation Network (HCGANet).

425, **TITLE:** Cost Affinity Learning Network for Stereo Matching

<https://doi.org/10.1109/ICASSP39728.2021.9413556>

**AUTHORS:** S. Chen, B. Li, W. Wang, H. Zhang, H. Li and Z. Wang

**HIGHLIGHT:** In this work, we propose a novel cost affinity learning network(CAL-Net) whose Affinity Enhanced Module(AEM) extracts the affinity of the elements in the cost feature and reconstructs a more discriminative feature.

426, **TITLE:** Video Quality Prediction Using Voxel-Wise fMRI Models of the Visual Cortex

<https://doi.org/10.1109/ICASSP39728.2021.9414328>

**AUTHORS:** N. S. Mahankali and S. S. Channappayya

**HIGHLIGHT:** In this work, we address the problem of full-reference video quality prediction.

427, **TITLE:** Tensor Decomposition Via Core Tensor Networks

<https://doi.org/10.1109/ICASSP39728.2021.9413637>

**AUTHORS:** J. ZHANG, Z. TAO, L. ZHANG and Q. ZHAO

**HIGHLIGHT:** In this paper, we propose an efficient TD algorithm that aims to learn a global mapping from input tensors to latent core tensors, under the assumption that the mappings of multiple tensors might be shared or highly correlated.

428, **TITLE:** Sign Language Segmentation with Temporal Convolutional Networks

<https://doi.org/10.1109/ICASSP39728.2021.9413817>

**AUTHORS:** K. Renz, N. C. Stache, S. Albanie and G. Varol

**HIGHLIGHT:** The objective of this work is to determine the location of temporal boundaries between signs in continuous sign language videos.

429, **TITLE:** An Adaptive Discriminant and Sparsity Feature Descriptor for Finger Vein Recognition

<https://doi.org/10.1109/ICASSP39728.2021.9413688>

**AUTHORS:** S. Li and B. Zhang

**HIGHLIGHT:** In this paper, we propose an adaptive discriminant and sparsity feature descriptor (DSFD) for FV feature extraction and recognition.

430, **TITLE:** Routinggan: Routing Age Progression and Regression with Disentangled Learning

<https://doi.org/10.1109/ICASSP39728.2021.9414735>

**AUTHORS:** Z. Huang, J. Zhang and H. Shan

**HIGHLIGHT:** To address these deficiencies and have the best of both worlds, this paper introduces a dropout-like method based on GAN (RoutingGAN) to route different effects in a high-level semantic feature space.

431, **TITLE:** Semantic-Aware Unpaired Image-to-Image Translation for Urban Scene Images

<https://doi.org/10.1109/ICASSP39728.2021.9414192>

**AUTHORS:** Z. Li, R. Togo, T. Ogawa and M. Haseyama

**HIGHLIGHT:** To tackle this problem, in this paper, we reasonably modify the previous problem setting and present a novel semantic-aware method.

432, **TITLE:** Fontnet: On-Device Font Understanding and Prediction Pipeline

<https://doi.org/10.1109/ICASSP39728.2021.9413410>

**AUTHORS:** R. S, R. Khurana, V. Agarwal, J. R. Vachhani and G. Bhanodai

**HIGHLIGHT:** In this paper, we propose two engines: Font Detection Engine, which identifies the font style, color and size attributes of text in an image and a Font Prediction Engine, which predicts similar fonts for a query font.

- 433, TITLE: Agent-Environment Network for Temporal Action Proposal Generation  
<https://doi.org/10.1109/ICASSP39728.2021.9415101>  
AUTHORS: V. -K. Vo-Ho, N. Le, K. Kamazaki, A. Sugimoto and M. -T. Tran  
HIGHLIGHT: Based on the action definition that a human, known as an agent, interacts with the environment and performs an action that affects the environment, we propose a contextual Agent-Environment Network.
- 434, TITLE: Adaptive Multi-Domain Learning for Outdoor 3d Human Pose and Shape Estimation  
<https://doi.org/10.1109/ICASSP39728.2021.9414590>  
AUTHORS: Z. Gui, S. Zhang, K. Wang, J. Yang and P. C. Yuen  
HIGHLIGHT: In this paper, we first point out this problem and then address it via a novel cascade multi-domain learning module (CMDL), where multiple adapters are employed to extract more discriminative features for different domains.
- 435, TITLE: Lightweight Human Pose Estimation under Resource-Limited Scenes  
<https://doi.org/10.1109/ICASSP39728.2021.9413748>  
AUTHORS: Z. Zhang, J. Tang and G. Wu  
HIGHLIGHT: In this paper, we investigate the problem of lightweight human pose estimation under resource-limited scenes. We first redesign a lightweight bottleneck block with two concepts: depthwise convolution and attention mechanism.
- 436, TITLE: Absolute 3d Pose Estimation and Length Measurement of Severely Deformed Fish from Monocular Videos in Longline Fishing  
<https://doi.org/10.1109/ICASSP39728.2021.9414803>  
AUTHORS: J. Mei, J. -N. Hwang, S. Romain, C. Rose, B. Moore and K. Magrane  
HIGHLIGHT: Unlike related works, which either require expensive 3D ground-truth data and/or multiple-view images to provide depth information, or are limited to rigid objects, we propose a novel frame-based method to estimate the absolute 3D fish pose and fish length from a single-view 2D segmentation mask.
- 437, TITLE: Camera Calibration with Pose Guidance  
<https://doi.org/10.1109/ICASSP39728.2021.9414468>  
AUTHORS: Y. Ren and F. Hu  
HIGHLIGHT: To resolve above issues, we propose a calibration system called Calibration with Pose Guidance to improve calibration accuracy, reduce calibration variance among different users or different trials of the same person.
- 438, TITLE: Real Versus Fake 4k - Authentic Resolution Assessment  
<https://doi.org/10.1109/ICASSP39728.2021.9413898>  
AUTHORS: R. R. Shah, V. Anirudh Akundy and Z. Wang  
HIGHLIGHT: This work aims at authentic resolution assessment (ARA).
- 439, TITLE: Perceptual Quality Assessment for Recognizing True and Pseudo 4k Content  
<https://doi.org/10.1109/ICASSP39728.2021.9414932>  
AUTHORS: W. Zhu, G. Zhai, X. Min, X. Yang and X. -P. Zhang  
HIGHLIGHT: To meet the imperative demand for monitoring the quality of Ultra High-Definition (UHD) content in multimedia industries, we propose an efficient no-reference (NR) image quality assessment (IQA) metric to distinguish original and pseudo 4K contents and measure the quality of their quality in this paper.
- 440, TITLE: A New Tubular Structure Tracking Algorithm Based On Curvature-Penalized Perceptual Grouping  
<https://doi.org/10.1109/ICASSP39728.2021.9414114>  
AUTHORS: L. Liu, D. Chen, M. Shu, H. Shu and L. D. Cohen  
HIGHLIGHT: In this paper, we propose a new minimal path-based framework for minimally interactive tubular structure tracking in conjunction with a perceptual grouping scheme.
- 441, TITLE: Multiple Human Tracking in Non-Specific Coverage with Wearable Cameras  
<https://doi.org/10.1109/ICASSP39728.2021.9414747>  
AUTHORS: S. Wang, R. Han, W. Feng and S. Wang  
HIGHLIGHT: To address this problem, in this paper we propose a Markov Decision Process with jump state (JMDP) to model the target's lifetime in tracking, and use optical flow of the camera motion and the statistical information of the targets to model the camera state transition.
- 442, TITLE: Fine-Grained Pose Temporal Memory Module for Video Pose Estimation and Tracking  
<https://doi.org/10.1109/ICASSP39728.2021.9413650>

AUTHORS: C. WANG et al.  
HIGHLIGHT: To better solve these problems and utilize the temporal information efficiently and effectively, we present a novel structure, called pose temporal memory module, which is flexible to be transferred into top-down pose estimation frameworks.

443, TITLE: Drawing Order Recovery from Trajectory Components  
<https://doi.org/10.1109/ICASSP39728.2021.9413542>  
AUTHORS: M. Yang, X. Zhou, Y. Sun, J. Chen and B. Qiang  
HIGHLIGHT: Based on the idea that drawing trajectories are able to be recovered by connecting their trajectory components in correct orders, this work proposes a novel DOR method from static images.

444, TITLE: Deep Hashing for Motion Capture Data Retrieval  
<https://doi.org/10.1109/ICASSP39728.2021.9413505>  
AUTHORS: N. Lv, Y. Wang, Z. Feng and J. Peng  
HIGHLIGHT: In this work, we propose an efficient retrieval method for human motion capture (MoCap) data based on supervised deep hash code learning.

445, TITLE: Hierarchical Attention Fusion for Geo-Localization  
<https://doi.org/10.1109/ICASSP39728.2021.9414517>  
AUTHORS: L. Yan, Y. Cui, Y. Chen and D. Liu  
HIGHLIGHT: In this work, we cast the geo-localization as a 2D image retrieval task.

446, TITLE: AttentionLite: Towards Efficient Self-Attention Models for Vision  
<https://doi.org/10.1109/ICASSP39728.2021.9415117>  
AUTHORS: S. Kundu and S. Sundaresan  
HIGHLIGHT: We propose a novel framework for producing a class of parameter and compute efficient models called AttentionLite suitable for resource constrained applications.

447, TITLE: Attention-Guided Second-Order Pooling Convolutional Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414964>  
AUTHORS: S. Chen, Q. Sun, C. Li, J. Zhang and Q. Zhang  
HIGHLIGHT: To handle above limitations, this paper proposes a novel attention-guided second-order pooling convolutional network (ASP-Net).

448, TITLE: SA-Net: Shuffle Attention for Deep Convolutional Neural Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414568>  
AUTHORS: Q. -L. Zhang and Y. -B. Yang  
HIGHLIGHT: In this paper, we propose an efficient Shuffle Attention (SA) module to address this issue, which adopts Shuffle Units to combine two types of attention mechanisms effectively.

449, TITLE: An Attention Based Wavelet Convolutional Model for Visual Saliency Detection  
<https://doi.org/10.1109/ICASSP39728.2021.9413755>  
AUTHORS: R. S. Bhooshan and S. K  
HIGHLIGHT: In this work, an attention based Wavelet Convolutional Neural Network (WCNN) is implemented that efficiently extracts the spatial, spectral and semantic features of the image in multiple resolution and it turns out to be suitable for locating the visually salient regions.

450, TITLE: Cascade Attention Fusion for Fine-Grained Image Captioning Based on Multi-Layer LSTM  
<https://doi.org/10.1109/ICASSP39728.2021.9413691>  
AUTHORS: S. Wang et al.  
HIGHLIGHT: In this paper, we propose a visual and semantic fusion network with a margin-based training guidance mechanism to generate fine image descriptions that depict more objects, attributes and other distinguishing aspects of images.

451, TITLE: Webly Supervised Deep Attentive Quantization  
<https://doi.org/10.1109/ICASSP39728.2021.9414172>  
AUTHORS: J. Wang, B. Chen, T. Dai and S. -T. Xia  
HIGHLIGHT: To solve this problem, we propose a novel method termed Webly Supervised Deep Attentive Quantization (WSDAQ), where deep quantization is trained on web images associated with some userprovided weak tags, without consulting any ground-truth labels.

452, TITLE: Unsupervised Audio-Visual Subspace Alignment for High-Stakes Deception Detection

<https://doi.org/10.1109/ICASSP39728.2021.9413550>

AUTHORS: L. MATHUR and M. J. MATARIC

HIGHLIGHT: To address this problem, we propose the first multimodal unsupervised transfer learning approach that detects real-world, high-stakes deception in videos with-out using high-stakes labels.

453, TITLE: Violence Detection in Videos Based on Fusing Visual and Audio Information

<https://doi.org/10.1109/ICASSP39728.2021.9413686>

AUTHORS: W. -F. Pang, Q. -H. He, Y. -j. Hu and Y. -X. Li

HIGHLIGHT: We proposed a neural network containing three modules for fusing multimodal information: 1) attention module for utilizing weighted features to generate effective features based on the mutual guidance between visual and audio information; 2) fusion module for integrating features by fusing visual and audio information based on the bilinear pooling mechanism; and 3) mutual Learning module for enabling the model to learn visual information from another neural network with a different architecture.

454, TITLE: QUERYD: A Video Dataset with High-Quality Text and Audio Narrations

<https://doi.org/10.1109/ICASSP39728.2021.9414640>

AUTHORS: A. -M. Oncescu, J. F. Henriques, Y. Liu, A. Zisserman and S. Albanie

HIGHLIGHT: We introduce QuerYD, a new large-scale dataset for retrieval and event localisation in video.

455, TITLE: Generating Natural Questions from Images for Multimodal Assistants

<https://doi.org/10.1109/ICASSP39728.2021.9413599>

AUTHORS: A. Patel, A. Bindal, H. Kotek, C. Klein and J. Williams

HIGHLIGHT: In this paper, we present an approach for generating diverse and meaningful questions that consider image content and metadata of image (e.g., location, associated keyword).

456, TITLE: An Adaptive Multi-Scale and Multi-Level Features Fusion Network with Perceptual Loss for Change Detection

<https://doi.org/10.1109/ICASSP39728.2021.9414394>

AUTHORS: J. Xu, Y. Luo, X. Chen and C. Luo

HIGHLIGHT: This paper proposes a novel adaptive multi-scale and multi-level features fusion network for change detection in very-high-resolution bi-temporal remote sensing images.

457, TITLE: SeeHear: Signer Diarisation and a New Dataset

<https://doi.org/10.1109/ICASSP39728.2021.9414856>

AUTHORS: S. Albanie et al.

HIGHLIGHT: In this work, we propose a framework to collect a large-scale, diverse sign language dataset that can be used to train automatic sign language recognition models. The first contribution of this work is SDTrack, a generic method for signer tracking and diarisation in the wild.

458, TITLE: Semantic Image Synthesis from Inaccurate and Coarse Masks

<https://doi.org/10.1109/ICASSP39728.2021.9414521>

AUTHORS: K. Katsumata and H. Nakayama

HIGHLIGHT: To address this issue, we propose a smoothing method, which we call local label smoothing (LLS), that incorporates label smoothing per small patch of an input mask to learn mapping from masks to images even when semantic masks are inaccurate.

459, TITLE: Range Guided Depth Refinement and Uncertainty-Aware Aggregation for View Synthesis

<https://doi.org/10.1109/ICASSP39728.2021.9413981>

AUTHORS: Y. Chang, Y. Chen and G. Wang

HIGHLIGHT: In this paper, we present a framework of view synthesis, including range guided depth refinement and uncertainty-aware aggregation based novel view synthesis.

460, TITLE: DP-VTON: Toward Detail-Preserving Image-Based Virtual Try-on Network

<https://doi.org/10.1109/ICASSP39728.2021.9414874>

AUTHORS: Y. Chang et al.

HIGHLIGHT: To resolve this issue, we present a novel virtual try-on network, DP-VTON.

461, TITLE: Light Field Style Transfer with Local Angular Consistency

<https://doi.org/10.1109/ICASSP39728.2021.9414689>

AUTHORS: D. Egan, M. Alain and A. Smolic

HIGHLIGHT: In this paper, we present a novel optimization-based method for light field style transfer which iteratively propagates the style from the centre view towards the outer views while enforcing local angular consistency.

- 462, TITLE: Skip Attention GAN for Remote Sensing Image Synthesis  
<https://doi.org/10.1109/ICASSP39728.2021.9414701>  
AUTHORS: K. Deng, K. Zhang, P. Yao, S. Cheng and P. He  
HIGHLIGHT: We establish Skip Attention Mechanism to deal with this problem, which learns dependencies between local points on low-resolution feature maps, and then upsample the attention map and combine it with high-resolution feature maps.
- 463, TITLE: Image Generation Based on Texture Guided VAE-AGAN for Regions of Interest Detection in Remote Sensing Images  
<https://doi.org/10.1109/ICASSP39728.2021.9413823>  
AUTHORS: L. Zhang and Y. Liu  
HIGHLIGHT: To cope with this issue, we propose a novel method based on texture guided variational autoencoder-attention wise generative adversarial network (VAE-AGAN) to augment the training data for ROI detection.
- 464, TITLE: EADNet: Efficient Asymmetric Dilated Network For Semantic Segmentation  
<https://doi.org/10.1109/ICASSP39728.2021.9413767>  
AUTHORS: Q. Yang, T. Chen, J. Fan, Y. Lu, C. Zuo and Q. Chi  
HIGHLIGHT: In this paper, we propose an efficient asymmetric dilated semantic segmentation network, named EADNet, which consists of multiple developed asymmetric convolution branches with different dilation rates to capture the variable shapes and scales information of an image.
- 465, TITLE: Ltaf-Net: Learning Task-Aware Adaptive Features and Refining Mask for Few-Shot Semantic Segmentation  
<https://doi.org/10.1109/ICASSP39728.2021.9414786>  
AUTHORS: B. Mao, L. Wang, S. Xiang and C. Pan  
HIGHLIGHT: In this paper we propose a novel model named LTAF-Net for few-shot segmentation.
- 466, TITLE: Cgan-Net: Class-Guided Asymmetric Non-Local Network for Real-Time Semantic Segmentation  
<https://doi.org/10.1109/ICASSP39728.2021.9414957>  
AUTHORS: H. Chen, Q. Hu, J. Yang, J. Wu and Y. Guo  
HIGHLIGHT: In this paper, we introduce a Class-Guided Asymmetric Non-local Network (CGAN-Net) to enhance the class-discriminability in learned feature map, while maintaining real-time efficiency.
- 467, TITLE: Aggregation Architecture and all-to-one Network for Real-Time Semantic Segmentation  
<https://doi.org/10.1109/ICASSP39728.2021.9413454>  
AUTHORS: K. CAO, X. HUANG and J. SHAO  
HIGHLIGHT: In this paper, we make the following contributions: (i) First, unlike the previous three architectures, we propose a new aggregation architecture as the network back-bone. (ii) Second, a multi-level auxiliary loss design model is used for the training phase, which can improve the model segmentation effect. (iii) According to this aggregation structure, an all-to-one network (ATONet) for real-time semantic segmentation is proposed, which achieves a good trade-off between speed and accuracy by assembling the features of all blocks.
- 468, TITLE: Nlkd: Using Coarse Annotations For Semantic Segmentation Based on Knowledge Distillation  
<https://doi.org/10.1109/ICASSP39728.2021.9414355>  
AUTHORS: D. Liang, Y. Du, H. Sun, L. Zhang, N. Liu and M. Wei  
HIGHLIGHT: This paper proposes a noise learning framework based on knowledge distillation NLKD, to improve segmentation performance on unclean data.
- 469, TITLE: Knowledge Reasoning for Semantic Segmentation  
<https://doi.org/10.1109/ICASSP39728.2021.9415022>  
AUTHORS: S. Chen, Z. Li and X. Yang  
HIGHLIGHT: To overcome the limitation of the traditional method, we propose a Knowledge Reasoning Net (KRNet) that consists of two crucial modules: (1) a prior knowledge mapping module that incorporates external knowledge by graph convolutional network to guide learning semantic representations and (2) a knowledge reasoning module that correlates these representations with a graph built on the external knowledge and explores their interactions via the knowledge reasoning.
- 470, TITLE: Non-Convex Sparse Deviation Modeling Via Generative Models  
<https://doi.org/10.1109/ICASSP39728.2021.9414170>  
AUTHORS: Y. Yang, H. Wang, H. Qiu, J. Wang and Y. Wang  
HIGHLIGHT: In this paper, the generative model is used to introduce the structural properties of the signal to replace the common sparse hypothesis, and a non-convex compressed sensing sparse deviation model based on the generative model (lq-Gen) is proposed.

- 471, TITLE: Imrnet: An Iterative Motion Compensation and Residual Reconstruction Network for Video Compressed Sensing  
<https://doi.org/10.1109/ICASSP39728.2021.9414534>  
AUTHORS: X. Yang and C. Yang  
HIGHLIGHT: This paper proposes an iterative motion compensation and residual reconstruction network for VCS, called ImrNet.
- 472, TITLE: Deep Color Constancy Using Temporal Gradient Under Ac Light Sources  
<https://doi.org/10.1109/ICASSP39728.2021.9413356>  
AUTHORS: J. -W. HA, J. -S. YOO and J. -O. KIM  
HIGHLIGHT: While most of conventional methods focus on only spatial information of a single image, we propose a deep spatio-temporal color constancy method.
- 473, TITLE: End-to-End Learning of Variational Models and Solvers for the Resolution of Interpolation Problems  
<https://doi.org/10.1109/ICASSP39728.2021.9414629>  
AUTHORS: R. Fablet, L. Drumetz and F. Rousseau  
HIGHLIGHT: We consider an application to inverse problems with incomplete datasets (image inpainting and multivariate time series interpolation).
- 474, TITLE: Multi-Models Fusion for Light Field Angular Super-Resolution  
<https://doi.org/10.1109/ICASSP39728.2021.9413824>  
AUTHORS: F. Cao, P. An, X. Huang, C. Yang and Q. Wu  
HIGHLIGHT: In this paper, therefore, we propose a multi-models fusion for LF SR in angular domain.
- 475, TITLE: Hide Chopin in the Music: Efficient Information Steganography Via Random Shuffling  
<https://doi.org/10.1109/ICASSP39728.2021.9413357>  
AUTHORS: Z. Sun, C. Li and Q. Zhao  
HIGHLIGHT: In this paper, we explore the room introduced by the low-rank property of natural signals (i.e., images, audios), and propose a training-free model for efficient information steganography, which provides a capacity of hiding full-size images into carriers of the same spatial resolution.
- 476, TITLE: Pointer Networks for Arbitrary-Shaped Text Spotting  
<https://doi.org/10.1109/ICASSP39728.2021.9414739>  
AUTHORS: Y. Zhang, W. Yang, Z. Xu, Y. Li, Z. Chen and L. Huang  
HIGHLIGHT: In this paper, we present a highly efficient one-stage method named PointerNet for arbitrary-shaped text spotting.
- 477, TITLE: Rotation Invariance Analysis of Local Convolutional Features in Image Retrieval  
<https://doi.org/10.1109/ICASSP39728.2021.9414491>  
AUTHORS: L. Zhao, Y. Wang and J. Kato  
HIGHLIGHT: In this work, our objective is to enhance the robustness of LC features against image rotation.
- 478, TITLE: Signature Feature Marking Enhanced IRM Framework for Drone Image Analysis in Precision Agriculture  
<https://doi.org/10.1109/ICASSP39728.2021.9413577>  
AUTHORS: A. Kadethankar, N. Sinha, V. Hegde and A. Burman  
HIGHLIGHT: In this work, we are proposing enhancement to Invariant Risk Minimization (IRM) framework which is Signature Feature Marking (SFM) enhanced IRM for object classification.
- 479, TITLE: Vehicle 3d Localization in Road Scenes VIA a Monocular Moving Camera  
<https://doi.org/10.1109/ICASSP39728.2021.9413487>  
AUTHORS: Y. Zhang, A. Zheng, K. Han, Y. Wang and J. -N. Hwang  
HIGHLIGHT: In this paper, we present an effective vehicle 3D localization method, that utilizes 2D key-points predicted from a trained CNN to model the vehicles' structure, from which the ground points are further inferred.
- 480, TITLE: Gps-Denied Navigation Using Sar Images And Neural Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414421>  
AUTHORS: T. White, J. Wheeler, C. Lindstrom, R. Christensen and K. R. Moon  
HIGHLIGHT: This paper presents a method for determining the navigation errors present at the beginning of a GPS-denied period utilizing data from a synthetic aperture radar (SAR) system.

481, TITLE: Attention-Embedded Decomposed Network with Unpaired CT Images Prior for Metal Artifact Reduction  
<https://doi.org/10.1109/ICASSP39728.2021.9413578>  
AUTHORS: B. Zhao, Q. Ren, J. Li and Y. Zhao  
HIGHLIGHT: To address these issues, we propose an attention-embedded decomposed network to reducing metal artifacts in both image space and sinogram space with unpaired images.

482, TITLE: Partial Feature Aggregation Network for Real-Time Object Counting  
<https://doi.org/10.1109/ICASSP39728.2021.9413835>  
AUTHORS: H. Yu and L. Zhang  
HIGHLIGHT: In this paper, we propose an efficient and accurate light-weight network for object counting, called Partial Feature Aggregation Network (PFANet).

483, TITLE: A Bayesian Inference Approach for Location-Based Micro Motions using Radio Frequency Sensing  
<https://doi.org/10.1109/ICASSP39728.2021.9414167>  
AUTHORS: D. A. Maluf, A. Elnakeeb and M. Silverman  
HIGHLIGHT: The target tracking objective is formulated as an inference problem, by which we show how the Bayesian framework can be exploited to infer the parameters of interest for a given physics model.

484, TITLE: Robust Spatial-Temporal Correlation Model for Background Initialization in Severe Scene  
<https://doi.org/10.1109/ICASSP39728.2021.9414796>  
AUTHORS: Y. Deng, W. Zhou, B. Peng, D. Liang and S. Kaneko  
HIGHLIGHT: In this work, we develop a novel method called co-occurrence pixel-block (CPB) model via spatial-temporal correlation for robust back-ground initialization.

485, TITLE: 2D-FRFT Based Frequency Shift-Invariant Digital Image Encryption  
<https://doi.org/10.1109/ICASSP39728.2021.9414984>  
AUTHORS: L. Gao, L. Qi and L. Guan  
HIGHLIGHT: In this paper, we study the property of frequency shift in two-dimensional Fractional Fourier Transform (2D-FRFT) domain.

486, TITLE: Capturing Banding in Images: Database Construction and Objective Assessment  
<https://doi.org/10.1109/ICASSP39728.2021.9413884>  
AUTHORS: A. Kapoor, J. Sapra and Z. Wang  
HIGHLIGHT: This work targets at capturing and quantifying banding artifacts in images.

487, TITLE: On The Camera Position Dithering In Visual 3d Reconstruction  
<https://doi.org/10.1109/ICASSP39728.2021.9414917>  
AUTHORS: Q. An and Y. Shen  
HIGHLIGHT: In this paper, we adopt statistical analysis for camera dithering in visual 3D reconstruction and demonstrate the improvement of reconstruction accuracy brought by camera dithering under some conditions.

488, TITLE: Long-Short Temporal Modeling for Efficient Action Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9413807>  
AUTHORS: L. Wu, Y. Zou and C. Zhang  
HIGHLIGHT: In this paper, we propose a new two-stream action recognition network, termed as MENet, consisting of a Motion Enhancement (ME) module and a Video-level Aggregation (VLA) module to achieve long-short temporal modeling.

489, TITLE: Multi-Directional Convolution Networks with Spatial-Temporal Feature Pyramid Module for Action Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9413610>  
AUTHORS: B. Yang, Z. Wang, W. Ran, H. Lu and Y. -P. P. Chen  
HIGHLIGHT: In this paper, we propose a novel and effective Multi-Directional Convolution (MDConv), which extracts features along different spatial-temporal orientations.

490, TITLE: Unsupervised Motion Representation Enhanced Network for Action Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9414222>  
AUTHORS: X. Yang, L. Kong and J. Yang  
HIGHLIGHT: To fill the gap, we propose UF-TSN, a novel end-to-end action recognition approach enhanced with an embedded lightweight unsupervised optical flow estimator.



- 491, TITLE: An Improved Deep Relation Network for Action Recognition in Still Images  
<https://doi.org/10.1109/ICASSP39728.2021.9414302>  
AUTHORS: W. Wu and J. Yu  
HIGHLIGHT: To this end, we propose an efficient relation module that combines Human-Object and Scene-Object relations for action recognition.
- 492, TITLE: Human-Aware Coarse-to-Fine Online Action Detection  
<https://doi.org/10.1109/ICASSP39728.2021.9413368>  
AUTHORS: Z. Yang, D. Huang, J. Qin and Y. Wang  
HIGHLIGHT: In this work, we propose a two-stage framework to efficiently and effectively detect actions on-the-fly.
- 493, TITLE: SRF-Net: Selective Receptive Field Network for Anchor-Free Temporal Action Detection  
<https://doi.org/10.1109/ICASSP39728.2021.9414253>  
AUTHORS: R. Ning, C. Zhang and Y. Zou  
HIGHLIGHT: In this study, we explore to remove the requirement of pre-defined anchors for TAD methods.
- 494, TITLE: Semantic-Aware Context Aggregation for Image Inpainting  
<https://doi.org/10.1109/ICASSP39728.2021.9414065>  
AUTHORS: Z. Huang, C. Qin, R. Liu, Z. Weng and Y. Zhu  
HIGHLIGHT: To handle this problem, we propose a novel semantic-aware context aggregation module (SACA) that aggregates distant contextual information from a semantic perspective by exploiting the internal semantic similarity of the input feature map.
- 495, TITLE: Bishift-Net for Image Inpainting  
<https://doi.org/10.1109/ICASSP39728.2021.9413380>  
AUTHORS: X. Zhou, T. Dai, Y. Jiang and S. -T. Xia  
HIGHLIGHT: Inspired by this, we propose a new inpainting model, which is called BiShift-Net.
- 496, TITLE: OAS-Net: Occlusion Aware Sampling Network for Accurate Optical Flow  
<https://doi.org/10.1109/ICASSP39728.2021.9413531>  
AUTHORS: L. Kong, X. Yang and J. Yang  
HIGHLIGHT: To deal with these challenges, we propose a lightweight yet efficient optical flow network, named OAS-Net (occlusion aware sampling network) for accurate optical flow.
- 497, TITLE: Mask4D: 4D Convolution Network for Light Field Occlusion Removal  
<https://doi.org/10.1109/ICASSP39728.2021.9413449>  
AUTHORS: Y. Li et al.  
HIGHLIGHT: In this paper, we present a simple yet effective LF occlusion removal method name Mask4D, which is a 4D convolution-based encoder-decoder network.
- 498, TITLE: Self-Supervised Depth Estimation Via Implicit Cues from Videos  
<https://doi.org/10.1109/ICASSP39728.2021.9413407>  
AUTHORS: J. Wang, G. Zhang, Z. Wu, X. Li and L. Liu  
HIGHLIGHT: In this work, we improve the self-supervised learning framework for depth estimation using consecutive frames from monocular and stereo videos.
- 499, TITLE: Scene Completeness-Aware Lidar Depth Completion for Driving Scenario  
<https://doi.org/10.1109/ICASSP39728.2021.9414295>  
AUTHORS: C. -Y. Wu and U. Neumann  
HIGHLIGHT: This paper introduces Scene Completeness-Aware Depth Completion (SCADC) to complete raw lidar scans into dense depth maps with fine and complete scene structures.
- 500, TITLE: Semi-Supervised Feature Embedding for Data Sanitization in Real-World Events  
<https://doi.org/10.1109/ICASSP39728.2021.9414461>  
AUTHORS: B. Lavi, J. Nascimento and A. Rocha  
HIGHLIGHT: We address the issue of establishing which images represent an event of interest through a semi-supervised learning technique.
- 501, TITLE: Exposing GAN-Generated Faces Using Inconsistent Corneal Specular Highlights

<https://doi.org/10.1109/ICASSP39728.2021.9414582>

AUTHORS: S. Hu, Y. Li and S. Lyu

HIGHLIGHT: In this work, we show that GAN synthesized faces can be exposed with the inconsistent corneal specular highlights between two eyes.

502, TITLE: A Features Decoupling Method for Multiple Manipulations Identification in Image Operation Chains

<https://doi.org/10.1109/ICASSP39728.2021.9414628>

AUTHORS: J. Chen, X. Liao, W. Wang and Z. Qin

HIGHLIGHT: In this paper, we focus on identifying the manipulations in an image operation chain composed of multiple manipulations in a certain order.

503, TITLE: Subjective and Objective Evaluation of Deepfake Videos

<https://doi.org/10.1109/ICASSP39728.2021.9414258>

AUTHORS: P. Korshunov and S. Marcel

HIGHLIGHT: Therefore, this paper, presents a subjective study, which, using 60 naïve subjects, evaluates how hard it is for humans to see if a video is a deepfake or not.

504, TITLE: Forensicability of Deep Neural Network Inference Pipelines

<https://doi.org/10.1109/ICASSP39728.2021.9414301>

AUTHORS: A. Schlögl, T. Kupek and R. Böhme

HIGHLIGHT: We propose methods to infer properties of the execution environment of machine learning pipelines by tracing characteristic numerical deviations in observable outputs.

505, TITLE: SERN: Stance Extraction and Reasoning Network for Fake News Detection

<https://doi.org/10.1109/ICASSP39728.2021.9414787>

AUTHORS: J. Xie, S. Liu, R. Liu, Y. Zhang and Y. Zhu

HIGHLIGHT: To solve this problem, a novel Stance Extraction and Reasoning Network (SERN) is proposed to extract the stances implied in post-reply pairs implicitly and integrate the stance representations for fake news detection without manually labeling stances, which saves much time and effort.

506, TITLE: An Efficient Paper Anti-Counterfeiting Method Based on Microstructure Orientation Estimation

<https://doi.org/10.1109/ICASSP39728.2021.9415114>

AUTHORS: Y. Sun, X. Liao and J. Liu

HIGHLIGHT: In this paper, we focus on the efficient microstructure orientation estimation of paper surface for authentication.

507, TITLE: Learning Double-Compression Video Fingerprints Left From Social-Media Platforms

<https://doi.org/10.1109/ICASSP39728.2021.9413366>

AUTHORS: I. Amerini, A. Anagnostopoulos, L. Maiano and L. R. Celis

HIGHLIGHT: Most of the work performed so far on social media provenance has concentrated on images; in this paper, we propose a CNN architecture that analyzes video content to trace videos back to their social network of origin.

508, TITLE: Checking PRNU Usability on Modern Devices

<https://doi.org/10.1109/ICASSP39728.2021.9413611>

AUTHORS: C. Albisani, M. Iuliani and A. Piva

HIGHLIGHT: In this paper, we first highlight that wrong source attribution can happen on smartphones belonging to the same brand when images are acquired both in default and in bokeh mode.

509, TITLE: Handwritten Digits Reconstruction from Unlabelled Embeddings

<https://doi.org/10.1109/ICASSP39728.2021.9414885>

AUTHORS: T. Thebaud, G. Le Lan and A. Larcher

HIGHLIGHT: In this paper, we investigate template reconstruction attack of touchscreen biometrics, based on handwritten digits writer verification.

510, TITLE: Effect of Video Pixel-Binning on Source Attribution of Mixed Media

<https://doi.org/10.1109/ICASSP39728.2021.9415094>

AUTHORS: S. Taspinar, M. Mohanty and N. Memon

HIGHLIGHT: This paper investigates the attribution of visual media in the context of matching a video query object to an image fingerprint or vice versa.

511, TITLE: Combining Dynamic Image and Prediction Ensemble for Cross-Domain Face Anti-Spoofing

<https://doi.org/10.1109/ICASSP39728.2021.9413926>

AUTHORS: L. Lv et al.

HIGHLIGHT: Hence, we propose a source data-free domain adaptative face anti-spoofing framework to optimize the network in the target domain without using labeled source data via modeling it into a problem of learning with noisy labels.

512, TITLE: Label-Guided Dictionary Pair Learning for ECG Biometric Recognition

<https://doi.org/10.1109/ICASSP39728.2021.9413355>

AUTHORS: M. Ma, G. Yang, K. Wang, Y. Huang and Y. Yin

HIGHLIGHT: To overcome this limitation, our method, namely label-guided dictionary pair learning, aims to learn a projective dictionary and reconstructed dictionary jointly, which achieves signal representation and reconstruction simultaneously.

513, TITLE: Backdoor Attack Against Speaker Verification

<https://doi.org/10.1109/ICASSP39728.2021.9413468>

AUTHORS: T. Zhai, Y. Li, Z. Zhang, B. Wu, Y. Jiang and S. -T. Xia

HIGHLIGHT: In this paper, we demonstrate that it is possible to inject the hidden backdoor for infecting speaker verification models by poisoning the training data.

514, TITLE: Class-Conditional Defense GAN Against End-To-End Speech Attacks

<https://doi.org/10.1109/ICASSP39728.2021.9413626>

AUTHORS: M. Esmailpour, P. Cardinal and A. L. Koerich

HIGHLIGHT: In this paper we propose a novel defense approach against end-to-end adversarial attacks developed to fool advanced speech-to-text systems such as DeepSpeech and Lingvo.

515, TITLE: Selfgait: A Spatiotemporal Representation Learning Method for Self-Supervised Gait Recognition

<https://doi.org/10.1109/ICASSP39728.2021.9413894>

AUTHORS: Y. Liu, Y. Zeng, J. Pu, H. Shan, P. He and J. Zhang

HIGHLIGHT: In this work, we propose a self-supervised gait recognition method, termed SelfGait, which takes advantage of the massive, diverse, unlabeled gait data as a pre-training process to improve the representation abilities of spatiotemporal backbones.

516, TITLE: Attack on Practical Speaker Verification System Using Universal Adversarial Perturbations

<https://doi.org/10.1109/ICASSP39728.2021.9413467>

AUTHORS: W. Zhang et al.

HIGHLIGHT: A two-step algorithm is proposed to optimize the universal adversarial perturbation to be text-independent and has little effect on the authentication text recognition.

517, TITLE: Highly Efficient Protection of Biometric Face Samples with Selective JPEG2000 Encryption

<https://doi.org/10.1109/ICASSP39728.2021.9413941>

AUTHORS: H. Hofbauer, Y. Mart'nez-D'az, S. Kirchgasser, H. M'ndez-V'zquez and A. Uhl

HIGHLIGHT: In this paper we will show that selective encryption of face biometric samples is secure.

518, TITLE: Deep Auto-Encoding and Biohashing for Secure Finger Vein Recognition

<https://doi.org/10.1109/ICASSP39728.2021.9414498>

AUTHORS: H. O. Shahreza and S. Marcel

HIGHLIGHT: In this paper, we propose a deep-learning-based approach for secure finger vein recognition.

519, TITLE: Topic Sequence Embedding for User Identity Linkage from Heterogeneous Behavior Data

<https://doi.org/10.1109/ICASSP39728.2021.9415111>

AUTHORS: J. Yang, W. Zhou, W. Qian, J. Han and S. Hu

HIGHLIGHT: In this paper, a novel Topic Sequence Embedding (TSeqE) method is proposed to embed contextual information of topics to represent users' intrinsic characteristics for identity linkage.

520, TITLE: Looking Through Walls: Inferring Scenes from Video-Surveillance Encrypted Traffic

<https://doi.org/10.1109/ICASSP39728.2021.9414391>

AUTHORS: D. Mari et al.

HIGHLIGHT: In this paper, we show that it is possible to infer visual data by intercepting the encrypted video stream of a surveillance system, and how this may be leveraged to track the movements of a person inside the secured area.

521, TITLE: Optimal Attacking Strategy Against Online Reputation Systems with Consideration of the Message-Based Persuasion Phenomenon

<https://doi.org/10.1109/ICASSP39728.2021.9415012>

- AUTHORS: Z. Chen and H. V. Zhao  
HIGHLIGHT: This paper aims to analyze the optimal attacking strategy, especially when considering the "message-based persuasion" phenomenon where users' ratings tend to be influenced by earlier ones.
- 522, TITLE: STEP-GAN: A One-Class Anomaly Detection Model with Applications to Power System Security  
<https://doi.org/10.1109/ICASSP39728.2021.9415102>  
AUTHORS: M. Adiban, A. Safari and G. Salvi  
HIGHLIGHT: We propose a novel structure for the multi-generator generative adversarial network (GAN) to address the challenges of detecting adversarial attacks.
- 523, TITLE: Application-Layer DDOS Attacks with Multiple Emulation Dictionaries  
<https://doi.org/10.1109/ICASSP39728.2021.9413570>  
AUTHORS: M. Cirillo, M. D. Mauro, V. Matta and M. Tambasco  
HIGHLIGHT: We consider the problem of identifying the members of a botnet under an application-layer (L7) DDoS attack, where a target site is flooded with a large number of requests that emulate legitimate users' patterns.
- 524, TITLE: Secret Key Generation Over Wireless Channels using short Blocklength Multilevel Source Polar Coding  
<https://doi.org/10.1109/ICASSP39728.2021.9415041>  
AUTHORS: H. Hentil?, Y. Y. Shkel and V. Koivunen  
HIGHLIGHT: Inspired by the state-of-the-art performance provided by polar codes in the short blocklength regime for channel coding, we propose an explicit protocol based on polar codes for generating the secret keys.
- 525, TITLE: Efficient Network Protection Games Against Multiple Types Of Strategic Attackers  
<https://doi.org/10.1109/ICASSP39728.2021.9413758>  
AUTHORS: Z. Xu and M. Baykal-G?rsoy  
HIGHLIGHT: This paper considers network protection games against different types of attackers for a heterogeneous network system with N units.
- 526, TITLE: Detection Of Malicious DNS and Web Servers using Graph-Based Approaches  
<https://doi.org/10.1109/ICASSP39728.2021.9415092>  
AUTHORS: J. Jia, Z. Dong, J. Li and J. W. Stokes  
HIGHLIGHT: In this work, we propose both unsupervised and semi-supervised defenses based on the available knowledge of the defender.
- 527, TITLE: Low Complexity Secure P-Tensor Product Compressed Sensing Reconstruction Outsourcing and Identity Authentication in Cloud  
<https://doi.org/10.1109/ICASSP39728.2021.9413390>  
AUTHORS: M. Wang, D. Xiao and J. Liang  
HIGHLIGHT: Specifically, we propose a low complexity and secure PTP-CSR outsourcing model to protect the signal privacy, and further introduce user authentication and data verification services.
- 528, TITLE: Privacy-Preserving near Neighbor Search via Sparse Coding with Ambiguation  
<https://doi.org/10.1109/ICASSP39728.2021.9414115>  
AUTHORS: B. Razeghi, S. Ferdowsi, D. Kostadinov, F. P. Calmon and S. Voloshynovskiy  
HIGHLIGHT: In this paper, we propose a framework for privacy-preserving approximate near neighbor search via stochastic sparsifying encoding.
- 529, TITLE: Privacy-Preserving Optimal Insulin Dosing Decision  
<https://doi.org/10.1109/ICASSP39728.2021.9414807>  
AUTHORS: Z. Ying et al.  
HIGHLIGHT: In this paper, we propose a privacy-preserving optimal insulin dosing decision in the IoMT system (PIDM) to assist doctors in their decision-making with the patients privacy.
- 530, TITLE: Privacy-Accuracy Trade-Off of Inference as Service  
<https://doi.org/10.1109/ICASSP39728.2021.9413438>  
AUTHORS: Y. Jin and L. Lai  
HIGHLIGHT: In this paper, we propose a general framework to provide a desirable trade-off between inference accuracy and privacy protection in the inference as service scenario.
- 531, TITLE: Federated Learning with Local Differential Privacy: Trade-Offs Between Privacy, Utility, and Communication

<https://doi.org/10.1109/ICASSP39728.2021.9413764>

AUTHORS: M. Kim, O. G?nl? and R. F. Schaefer

HIGHLIGHT: We consider Gaussian mechanisms to preserve local differential privacy (LDP) of user data in the FL model with SGD.

532, TITLE: Scalable Privacy-Preserving Distributed Extremely Randomized Trees for Structured Data With Multiple Colluding Parties

<https://doi.org/10.1109/ICASSP39728.2021.9413632>

AUTHORS: A. Aminifar, F. Rabbi and Y. Lamo

HIGHLIGHT: In this paper, we extend the distributed Extremely Randomized Trees (ERT) approach w.r.t. privacy and scalability.

533, TITLE: Active Privacy-Utility Trade-Off Against A Hypothesis Testing Adversary

<https://doi.org/10.1109/ICASSP39728.2021.9414608>

AUTHORS: E. Erdemir, P. L. Dragotti and D. G?nd?z

HIGHLIGHT: We formulate both problems as a Markov decision process (MDP), and numerically solve them by advantage actor-critic (A2C) deep reinforcement learning (RL).

534, TITLE: Baitradar: A Multi-Model Clickbait Detection Algorithm Using Deep Learning

<https://doi.org/10.1109/ICASSP39728.2021.9414424>

AUTHORS: B. Gamage, A. Labib, A. Joomun, C. H. Lim and K. Wong

HIGHLIGHT: This issue is addressed in this study by proposing an algorithm called BaitRadard, which uses a deep learning technique where six inference models are jointly consulted to make the final classification decision.

535, TITLE: Enabling Efficient and Expressive Spatial Keyword Queries On Encrypted Data

<https://doi.org/10.1109/ICASSP39728.2021.9413414>

AUTHORS: X. Wang, J. Ma and X. Liu

HIGHLIGHT: In this paper, we propose a Secure Spatial Keyword Queries (SSKQ) construction supporting expressive query types.

536, TITLE: Privacy-Preserving Cloud-Based DNN Inference

<https://doi.org/10.1109/ICASSP39728.2021.9413820>

AUTHORS: S. Xie, B. Liu and Y. Hong

HIGHLIGHT: In this paper, we propose a novel privacy preserving cloud-based DNN inference framework ("PROUD"), which greatly improves the computational efficiency.

537, TITLE: Crypto-Oriented Neural Architecture Design

<https://doi.org/10.1109/ICASSP39728.2021.9413592>

AUTHORS: A. Shafraan, G. Segev, S. Peleg and Y. Hoshen

HIGHLIGHT: Differently, we propose to optimize the design of crypto-oriented neural architectures, introducing a novel Partial Activation layer.

538, TITLE: Integrating Deep Learning with First-Order Logic Programmed Constraints for Zero-Day Phishing Attack Detection

<https://doi.org/10.1109/ICASSP39728.2021.9414850>

AUTHORS: S. -J. Bu and S. -B. Cho

HIGHLIGHT: We introduce the integration method of deep learning and logic programmed domain knowledge to inject the real-world constraints.

539, TITLE: Improved Probabilistic Context-Free Grammars for Passwords Using Word Extraction

<https://doi.org/10.1109/ICASSP39728.2021.9414886>

AUTHORS: H. Cheng, W. Li, P. Wang and K. Liang

HIGHLIGHT: In this paper, we propose a word extraction approach for passwords, and further present an improved PCFG model, called WordPCFG.

540, TITLE: Enhancing Image Steganography Via Stego Generation And Selection

<https://doi.org/10.1109/ICASSP39728.2021.9414723>

AUTHORS: T. Song, M. Liu, W. Luo and P. Zheng

HIGHLIGHT: Unlike most existing steganography methods which are mainly focused on designing embedding cost, in this paper, we propose a new method to enhance existing steganographic methods via stego generation and selection.

- 541, TITLE: Synchronous Multi-Bit Audio Watermarking Based on Phase Shifting  
<https://doi.org/10.1109/ICASSP39728.2021.9414307>  
AUTHORS: S. Wang, W. Yuan, Z. Zhang, J. Wang and M. Unoki  
HIGHLIGHT: We considered the use of the distribution of the phase spectrum and propose an effective multi-bit audio watermarking method based on phase shifting.
- 542, TITLE: Image Steganography Based on Iterative Adversarial Perturbations Onto a Synchronized-Directions Sub-Image  
<https://doi.org/10.1109/ICASSP39728.2021.9414055>  
AUTHORS: X. Qin, S. Tan, W. Tang, B. Li and J. Huang  
HIGHLIGHT: In this paper, we present a novel steganographic scheme to incorporate synchronizing modification directions and iterative adversarial perturbations to enhance steganographic performance.
- 543, TITLE: Extending the Reverse JPEG Compatibility Attack to Double Compressed Images  
<https://doi.org/10.1109/ICASSP39728.2021.9414204>  
AUTHORS: J. Butora and J. Fridrich  
HIGHLIGHT: In this paper, we provide mathematical analysis and demonstrate experimentally that this attack can be extended to double compressed images when the first compression quality is 93 or larger and the second quality equal or larger than the first quality.
- 544, TITLE: Reversible Data Hiding in Jpeg Images for Privacy Protection  
<https://doi.org/10.1109/ICASSP39728.2021.9415118>  
AUTHORS: Y. Huang, X. Cao, H. -T. Wu and Y. -m. Cheung  
HIGHLIGHT: In this paper, an improved algorithm is proposed to conceal privacy information in JPEG images.
- 545, TITLE: A Layered Embedding-Based Scheme to Cope with Intra-Frame Distortion Drift In IPM-Based HEVC  
Steganography  
<https://doi.org/10.1109/ICASSP39728.2021.9413728>  
AUTHORS: X. Jia, J. Wang, Y. Liu, X. Kang and Y. -q. Shi  
HIGHLIGHT: To solve this problem, we propose a layered embedding scheme which embeds information into the intra-prediction modes (IPMs) of 4x4 intra-frame prediction units (PUs) in HEVC.
- 546, TITLE: Meta-Learning with Attention for Improved Few-Shot Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9414936>  
AUTHORS: Z. Hou, A. Walid and S. -Y. Kung  
HIGHLIGHT: In this paper, we propose meta-learning with attention mechanisms.
- 547, TITLE: B-Small: A Bayesian Neural Network Approach to Sparse Model-Agnostic Meta-Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9414437>  
AUTHORS: A. Madan and R. Prasad  
HIGHLIGHT: In this paper, we propose a Bayesian neural network based MAML algorithm, which we refer to as the B-SMALL algorithm.
- 548, TITLE: Deep Transform and Metric Learning Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414990>  
AUTHORS: W. Tang, E. Chouzenoux, J. -C. Pesquet and H. Krim  
HIGHLIGHT: We hence propose a novel Deep DL approach where each DL layer can be formulated and solved as a combination of one linear layer and a Recurrent Neural Network, where the RNN is flexibly regraded as a layer-associated learned metric.
- 549, TITLE: Robustness and Diversity Seeking Data-Free Knowledge Distillation  
<https://doi.org/10.1109/ICASSP39728.2021.9414674>  
AUTHORS: P. Han, J. Park, S. Wang and Y. Liu  
HIGHLIGHT: To overcome this challenge, we propose robustness and diversity seeking data-free KD (RDSKD) in this paper.
- 550, TITLE: Ensemble Distillation Approaches for Grammatical Error Correction  
<https://doi.org/10.1109/ICASSP39728.2021.9413385>  
AUTHORS: Y. Fathullah, M. J. F. Gales and A. Malinin  
HIGHLIGHT: This paper examines the application of both these distillation approaches to a sequence prediction task, grammatical error correction (GEC).

- 551, TITLE: Train Your Classifier First: Cascade Neural Networks Training from Upper Layers to Lower Layers  
<https://doi.org/10.1109/ICASSP39728.2021.9413565>  
AUTHORS: S. Zhang, C. -T. Do, R. Doddipatla, E. Loweimi, P. Bell and S. Renals  
HIGHLIGHT: In this paper, for the first time, we show that the frozen classifier is transferable within the same dataset.
- 552, TITLE: How Convolutional Neural Networks Deal with Aliasing  
<https://doi.org/10.1109/ICASSP39728.2021.9414627>  
AUTHORS: A. H. Ribeiro and T. B. Sch?n  
HIGHLIGHT: The question we aim to answer in this paper is simply: "how and to what extent do CNNs counteract aliasing?"
- 553, TITLE: Canet: Context-Aware Loss for Descriptor Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9415018>  
AUTHORS: T. Chen, X. Hu, J. Xiao, G. Zhang and H. Ruan  
HIGHLIGHT: In this paper, we propose a novel loss function to introduce more context information to facilitate training.
- 554, TITLE: Progressive Multi-Stage Feature Mix for Person Re-Identification  
<https://doi.org/10.1109/ICASSP39728.2021.9413533>  
AUTHORS: Y. Zhang, B. He, L. Sun and Q. Li  
HIGHLIGHT: In this work, we propose a Progressive Multi-stage feature Mix network (PMM), which enables the model to find out the more precise and diverse features in a progressive manner.
- 555, TITLE: Using Deep Image Priors to Generate Counterfactual Explanations  
<https://doi.org/10.1109/ICASSP39728.2021.9413636>  
AUTHORS: V. Narayanaswamy, J. J. Thiagarajan and A. Spanias  
HIGHLIGHT: To this end, we propose a novel regularization strategy based on an auxiliary loss estimator jointly trained with the predictor, which efficiently guides the prior to re-cover natural pre-images.
- 556, TITLE: Elliptical Shape Recovery from Blurred Pixels Using Deep Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9413440>  
AUTHORS: H. Zamani, P. Rostami, A. Amini and F. Marvasti  
HIGHLIGHT: In this paper, we study the problem of ellipse recovery from blurred shape images.
- 557, TITLE: Factorized CRF with Batch Normalization Based on the Entire Training Data  
<https://doi.org/10.1109/ICASSP39728.2021.9414188>  
AUTHORS: E. Goldman and J. Goldberger  
HIGHLIGHT: In this study we show that it is feasible to calculate the mean and variance using the entire training dataset instead of standard BN for any network node obtained as a linear function of the input features.
- 558, TITLE: Evolutionary Quantization of Neural Networks with Mixed-Precision  
<https://doi.org/10.1109/ICASSP39728.2021.9413631>  
AUTHORS: Z. Liu, X. Zhang, S. Wang, S. Ma and W. Gao  
HIGHLIGHT: In this paper, we develop a novel evolutionary based method to automatically determine the bit-widths of weights and activations in each convolutional layer, namely, Evolutionary Mixed-Precision Quantization (EMQ).
- 559, TITLE: Evolving Quantized Neural Networks for Image Classification Using A Multi-Objective Genetic Algorithm  
<https://doi.org/10.1109/ICASSP39728.2021.9413519>  
AUTHORS: Y. Wang, X. Wang and X. He  
HIGHLIGHT: To address this problem, this paper proposes a mixed precision quantization method combined with channel expansion of CNNs by using a multi-objective genetic algorithm, called MOGAQNN.
- 560, TITLE: Spectral Domain Convolutional Neural Network  
<https://doi.org/10.1109/ICASSP39728.2021.9413409>  
AUTHORS: B. Guan, J. Zhang, W. A. Sethares, R. Kijowski and F. Liu  
HIGHLIGHT: Although embedding CNN architectures in the spectral do-main is widely exploited to accelerate the training process, we demonstrate that it is also possible to use the spectral domain to reduce the memory footprint, a method we call Spectral Domain Convolutional Neural Network (SpecNet) that performs both the convolution and the activation operations in the spectral domain.
- 561, TITLE: Parametric Spectral Filters for Fast Converging, Scalable Convolutional Neural Networks



<https://doi.org/10.1109/ICASSP39728.2021.9414587>

AUTHORS: L. Wood and E. C. Larson

HIGHLIGHT: To address this, we propose the usage of spectral parametric functions to represent massive spectral domain filters with only a few trainable parameters.

562, TITLE: Feature Reuse for a Randomization Based Neural Network

<https://doi.org/10.1109/ICASSP39728.2021.9413424>

AUTHORS: X. Liang, M. Skoglund and S. Chatterjee

HIGHLIGHT: We propose a feature reuse approach for an existing multi-layer randomization based feedforward neural network.

563, TITLE: A ReLU Dense Layer to Improve the Performance of Neural Networks

<https://doi.org/10.1109/ICASSP39728.2021.9414269>

AUTHORS: A. M. Javid, S. Das, M. Skoglund and S. Chatterjee

HIGHLIGHT: We propose ReDense as a simple and low complexity way to improve the performance of trained neural networks.

564, TITLE: Nested Learning for Multi-Level Classification

<https://doi.org/10.1109/ICASSP39728.2021.9415076>

AUTHORS: R. Achddou, J. M. Di Martino and G. Sapiro

HIGHLIGHT: In this work, we address this problem in the context of nested learning.

565, TITLE: Cross-Modal Representation Reconstruction for Zero-Shot Classification

<https://doi.org/10.1109/ICASSP39728.2021.9413572>

AUTHORS: Y. Wang and S. Zhao

HIGHLIGHT: In this paper, we propose a Cross-Modal Representation Reconstruction (CM-RR) framework to bridge the semantic gap between visual features and semantic attributes, as well as introducing a novel regularizer for automatically feature selection.

566, TITLE: HIGCNN: Hierarchical Interleaved Group Convolutional Neural Networks for Point Clouds Analysis

<https://doi.org/10.1109/ICASSP39728.2021.9413879>

AUTHORS: J. Dang and J. Yang

HIGHLIGHT: In this paper, we present an efficient and lightweight neural network for point clouds analysis, named HIGCNN, which can achieve better performance but lower complexity compared to existing methods.

567, TITLE: AutoKWS: Keyword Spotting with Differentiable Architecture Search

<https://doi.org/10.1109/ICASSP39728.2021.9414848>

AUTHORS: B. Zhang, W. Li, Q. Li, W. Zhuang, X. Chu and Y. Wang

HIGHLIGHT: In this paper, we propose to leverage recent advances in differentiable neural architecture search to discover more efficient networks.

568, TITLE: Embedding Semantic Hierarchy in Discrete Optimal Transport for Risk Minimization

<https://doi.org/10.1109/ICASSP39728.2021.9414891>

AUTHORS: Y. Ge et al.

HIGHLIGHT: In this paper, we propose to incorporate the risk-aware inter-class correlation in a discrete optimal transport (DOT) training framework by configuring its ground distance matrix.

569, TITLE: Identifying Spammers to Boost Crowdsourced Classification

<https://doi.org/10.1109/ICASSP39728.2021.9414242>

AUTHORS: P. A. Traganitis and G. B. Giannakis

HIGHLIGHT: To curb their effect, a novel spectral algorithm for spammer detection that utilizes second-order statistics of annotators, is developed and preliminary results on synthetic and real data showcase the potential of this approach.

570, TITLE: A Rank-Constrained Clustering Algorithm with Adaptive Embedding

<https://doi.org/10.1109/ICASSP39728.2021.9414490>

AUTHORS: S. Pei, F. Nie, R. Wang and X. Li

HIGHLIGHT: In this article, a novel Rank-Constrained clustering algorithm with Adaptive Embedding called RCAE is proposed, where the spectral embedding and the clustering structure are learned simultaneously, so, the influence of noise on performance is greatly reduced.

- 571, TITLE: Towards Efficient Age Estimation by Embedding Potential Gender Features  
<https://doi.org/10.1109/ICASSP39728.2021.9414529>  
AUTHORS: Y. Deng, L. Fei, S. Teng, W. Zhang, D. Liu and Y. Hou  
HIGHLIGHT: In this paper, we propose a simplified deep learning network for age estimation by simultaneously learning aging and potential gender features.
- 572, TITLE: Adversarial Attacks on Coarse-to-Fine Classifiers  
<https://doi.org/10.1109/ICASSP39728.2021.9413743>  
AUTHORS: I. R. Alkhouri and G. K. Atia  
HIGHLIGHT: In this paper, we examine the susceptibility of coarse-to-fine hierarchical classifiers to such types of attacks.
- 573, TITLE: GDTW: A Novel Differentiable DTW Loss for Time Series Tasks  
<https://doi.org/10.1109/ICASSP39728.2021.9413895>  
AUTHORS: X. Liu, N. Li and S. -T. Xia  
HIGHLIGHT: Based on DTW, we propose a novel loss function for time series data called Gumbel-Softmin based fast DTW (GDTW).
- 574, TITLE: Hierarchical Recurrent Neural Network for Handwritten Strokes Classification  
<https://doi.org/10.1109/ICASSP39728.2021.9413412>  
AUTHORS: I. Degtyarenko et al.  
HIGHLIGHT: The paper presents an original solution to the online handwritten document processing in a free form, which is aimed at separating multi-class handwritten documents into texts, tables, formulas, drawings, etc.
- 575, TITLE: Robust Domain-Free Domain Generalization with Class-Aware Alignment  
<https://doi.org/10.1109/ICASSP39728.2021.9413872>  
AUTHORS: W. Zhang, M. Ragab and R. Sagarna  
HIGHLIGHT: In this paper, we propose DomainFree Domain Generalization (DFDG), a model-agnostic method to achieve better generalization performance on the unseen test domain without the need for source domain labels.
- 576, TITLE: One-Bit Compressed Sensing Using Untrained Network Prior  
<https://doi.org/10.1109/ICASSP39728.2021.9414581>  
AUTHORS: S. Kafle, G. Joseph and P. K. Varshney  
HIGHLIGHT: In this paper, we address the problem of one-bit compressed sensing using the data-driven deep learning approach.
- 577, TITLE: Deep Unfolding Network for Block-Sparse Signal Recovery  
<https://doi.org/10.1109/ICASSP39728.2021.9414163>  
AUTHORS: R. Fu, V. Monardo, T. Huang and Y. Liu  
HIGHLIGHT: In this paper, we put forward a block-sparse reconstruction network named Ada-BlockLISTA based on the concept of deep unfolding.
- 578, TITLE: REST: Robust IEarned Shrinkage-Thresholding Network Taming Inverse Problems with Model Mismatch  
<https://doi.org/10.1109/ICASSP39728.2021.9414141>  
AUTHORS: W. Pu, C. Zhou, Y. C. Eldar and M. R. D. Rodrigues  
HIGHLIGHT: We consider compressive sensing problems with model mismatch where one wishes to recover a sparse high-dimensional vector from low-dimensional observations subject to uncertainty in the measurement operator.
- 579, TITLE: Unfolding Neural Networks for Compressive Multichannel Blind Deconvolution  
<https://doi.org/10.1109/ICASSP39728.2021.9414443>  
AUTHORS: B. Tolooshams, S. Mulleti, D. Ba and Y. C. Eldar  
HIGHLIGHT: We propose a learned-structured unfolding neural network for the problem of compressive sparse multichannel blind-deconvolution.
- 580, TITLE: Sparsity Driven Latent Space Sampling for Generative Prior Based Compressive Sensing  
<https://doi.org/10.1109/ICASSP39728.2021.9413451>  
AUTHORS: V. Killedar, P. K. Pokala and C. Sekhar Seelamantula  
HIGHLIGHT: We propose a proximal meta-learning (PML) algorithm to enforce sparsity in the latent-space while training the generator.
- 581, TITLE: A Sparse Coding Approach to Automatic Diet Monitoring with Continuous Glucose Monitors

<https://doi.org/10.1109/ICASSP39728.2021.9414452>

AUTHORS: A. Das et al.

HIGHLIGHT: This article presents an approach to estimate dietary intake automatically by analyzing the post-prandial glucose response (PPGR) of a meal, as measured with continuous glucose monitors.

582, TITLE: Speeding Up of Kernel-Based Learning for High-Order Tensors

<https://doi.org/10.1109/ICASSP39728.2021.9413524>

AUTHORS: O. Karmouda, J. Boulanger and R. Boyer

HIGHLIGHT: In this work, we propose a fast Grassmannian kernel-based method for high-order tensor learning based on the equivalence between the Tucker and the tencortrain decompositions.

583, TITLE: A Fast Randomized Adaptive CP Decomposition For Streaming Tensors

<https://doi.org/10.1109/ICASSP39728.2021.9413554>

AUTHORS: L. T. Thanh, K. Abed-Meraim, N. L. Trung and A. Hafiane

HIGHLIGHT: In this paper, we introduce a fast adaptive algorithm for CAN- DECOMP/PARAFAC decomposition of streaming three-way tensors using randomized sketching techniques.

584, TITLE: Rank-Revealing Block-Term Decomposition for Tensor Completion

<https://doi.org/10.1109/ICASSP39728.2021.9415104>

AUTHORS: A. A. Rontogiannis, P. V. Giampouras and E. Kofidis

HIGHLIGHT: In this paper, BTD is employed for the completion of a tensor from its partially observed entries.

585, TITLE: Kernel Learning with Tensor Networks

<https://doi.org/10.1109/ICASSP39728.2021.9413826>

AUTHORS: K. Konstantinidis, S. Li and D. P. Mandic

HIGHLIGHT: To this end, we introduce a Tensor Network (TN) approach to learning kernel embeddings, with a TN serving to map the input to a low dimensional manifold, where a suitable base kernel function can be applied.

586, TITLE: Fiber-Sampled Stochastic Mirror Descent for Tensor Decomposition with  $\ell_1$ -Divergence

<https://doi.org/10.1109/ICASSP39728.2021.9413830>

AUTHORS: W. Pu, S. Ibrahim, X. Fu and M. Hong

HIGHLIGHT: In this paper, a unified stochastic mirror descent framework is developed for large-scale  $\ell_1$ -divergence CPD.

587, TITLE: Regularized Recovery by Multi-Order Partial Hypergraph Total Variation

<https://doi.org/10.1109/ICASSP39728.2021.9414277>

AUTHORS: R. Qu, J. He, H. Feng, C. Xu and B. Hu

HIGHLIGHT: In this work, we take this divergence into consideration, and propose a multi-order hypergraph Laplacian and the corresponding total variation.

588, TITLE: Learning Discriminative Features for Semi-Supervised Anomaly Detection

<https://doi.org/10.1109/ICASSP39728.2021.9414285>

AUTHORS: Z. Feng, J. Tang, Y. Dou and G. Wu

HIGHLIGHT: In this paper, we consolidate the model's discriminative power by introducing a transfer learning scheme to anomaly detection, thereby the model suffers less perturbation caused by pollution.

589, TITLE: RGLN: Robust Residual Graph Learning Networks via Similarity-Preserving Mapping on Graphs

<https://doi.org/10.1109/ICASSP39728.2021.9414792>

AUTHORS: J. Tang, X. Gao and W. Hu

HIGHLIGHT: To this end, we propose a residual graph learning paradigm to infer edge connectivities and weights in graphs, which is cast as distance metric learning under a low-rank assumption and a similarity-preserving regularization.

590, TITLE: Sequence-Level Self-Teaching Regularization

<https://doi.org/10.1109/ICASSP39728.2021.9413644>

AUTHORS: E. Sun, L. Lu, Z. Meng and Y. Gong

HIGHLIGHT: In this paper, we extend the previous approach and propose a sequence self-teaching network to regularize the sequence-level information in speech recognition.

591, TITLE: Wearing A Mask: Compressed Representations of Variable-Length Sequences Using Recurrent Neural Tangent Kernels

<https://doi.org/10.1109/ICASSP39728.2021.9413450>

AUTHORS: S. Alemohammad et al.  
HIGHLIGHT: To address this gap, we extend existing methods that rely on the use of kernels to variable-length sequences via use of the Recurrent Neural Tangent Kernel (RNTK).

592, TITLE: H-GPR: A Hybrid Strategy for Large-Scale Gaussian Process Regression  
<https://doi.org/10.1109/ICASSP39728.2021.9414474>  
AUTHORS: N. Li, Y. Gao, W. Li, Y. Jiang and S. -T. Xia  
HIGHLIGHT: This paper proposes a hybrid strategy called H-GPR to combine these two well-established approaches.

593, TITLE: Learning Optimal Lattice Codes for MIMO Communications  
<https://doi.org/10.1109/ICASSP39728.2021.9414237>  
AUTHORS: L. Amor's and M. Pitk?nen  
HIGHLIGHT: We propose a novel reinforcement learning approach to learning lattice codes for MIMO channels.

594, TITLE: A Bayesian Interpretation of the Light Gated Recurrent Unit  
<https://doi.org/10.1109/ICASSP39728.2021.9414259>  
AUTHORS: A. Bittar and P. N. Garner  
HIGHLIGHT: We derive a layerwise recurrence without the assumptions of previous work, and show that it leads to a standard recurrence with modest modifications to reflect use of log-probabilities.

595, TITLE: A Large-Dimensional Analysis of Symmetric SNE  
<https://doi.org/10.1109/ICASSP39728.2021.9413583>  
AUTHORS: C. S?jourm?, R. Couillet and P. Comon  
HIGHLIGHT: This work provides first answers by leveraging a large dimensional statistics approach, where the number  $n$  and dimension  $p$  of the large-dimensional data are of the same magnitude.

596, TITLE: A Dynamical Systems Perspective on Online Bayesian Nonparametric Estimators with Adaptive Hyperparameters  
<https://doi.org/10.1109/ICASSP39728.2021.9414378>  
AUTHORS: A. Koppel, A. S. Bedi and V. Krishnamurthy  
HIGHLIGHT: This paper presents and analyzes constant step size stochastic gradient algorithms in reproducing kernel Hilbert Space (RKHS), which encapsulates various adaptive nonlinear interpolation schemes.

597, TITLE: Online Multi-Hop Information Based Kernel Learning Over Graphs  
<https://doi.org/10.1109/ICASSP39728.2021.9414086>  
AUTHORS: Z. Zong and Y. Shen  
HIGHLIGHT: A multikernel-based approach is developed, which is capable of leveraging global network information, and scales well with network size as well.

598, TITLE: Sparsity in Max-Plus Algebra and Applications in Multivariate Convex Regression  
<https://doi.org/10.1109/ICASSP39728.2021.9414139>  
AUTHORS: N. Tsilivis, A. Tsiamis and P. Maragos  
HIGHLIGHT: In this paper, we study concepts of sparsity in the max-plus algebra and apply them to the problem of multivariate convex regression.

599, TITLE: Complex-Valued Vs. Real-Valued Neural Networks for Classification Perspectives: An Example on Non-Circular Data  
<https://doi.org/10.1109/ICASSP39728.2021.9413814>  
AUTHORS: J. A. Barrachina, C. Ren, C. Morisseau, G. Vieillard and J. . -P. Ovarlez  
HIGHLIGHT: Motivated by radar and especially Synthetic Aperture Radar (SAR) applications, we propose a statistical analysis of fully connected feed-forward neural networks performance in the cases where real and imaginary parts of the data are correlated through the non-circular property.

600, TITLE: High-Frequency Adversarial Defense for Speech and Audio  
<https://doi.org/10.1109/ICASSP39728.2021.9414525>  
AUTHORS: R. Olivier, B. Raj and M. Shah  
HIGHLIGHT: We explore a smoothing approach based on additive noise masking in priority high frequencies.

601, TITLE: Learning Separable Time-Frequency Filterbanks for Audio Classification  
<https://doi.org/10.1109/ICASSP39728.2021.9414916>

- AUTHORS: J. Pu, Y. Panagakis and M. Pantic  
HIGHLIGHT: In this paper, we aim to learn audio representations directly from raw audio, and at the same time mitigate its training burden by employing a light-weight architecture.
- 602, TITLE: Upsampling Artifacts in Neural Audio Synthesis  
<https://doi.org/10.1109/ICASSP39728.2021.9414913>  
AUTHORS: J. Pons, S. Pascual, G. Cengarle and J. Serr?  
HIGHLIGHT: Here, we address this gap by studying this problem from the audio signal processing perspective.
- 603, TITLE: Deep Convolutional and Recurrent Networks for Polyphonic Instrument Classification from Monophonic Raw Audio Waveforms  
<https://doi.org/10.1109/ICASSP39728.2021.9413479>  
AUTHORS: K. Avramidis, A. Kratimenos, C. Garoufis, A. Zlatintsi and P. Maragos  
HIGHLIGHT: In this paper, we attempt to recognize musical instruments in polyphonic audio by only feeding their raw waveforms into deep learning models.
- 604, TITLE: Learning Audio Embeddings with User Listening Data for Content-Based Music Recommendation  
<https://doi.org/10.1109/ICASSP39728.2021.9414458>  
AUTHORS: K. Chen, B. Liang, X. Ma and M. Gu  
HIGHLIGHT: To combat this problem, we first explore user listening history and demographics to construct a user embedding representing the user's music preference.
- 605, TITLE: Efficient Speech Emotion Recognition Using Multi-Scale CNN and Attention  
<https://doi.org/10.1109/ICASSP39728.2021.9414286>  
AUTHORS: Z. Peng, Y. Lu, S. Pan and Y. Liu  
HIGHLIGHT: In this paper, we propose a simple yet efficient neural network architecture to exploit both acoustic and lexical information from speech.
- 606, TITLE: Neural Audio Fingerprint for High-Specific Audio Retrieval Based on Contrastive Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9414337>  
AUTHORS: S. Chang et al.  
HIGHLIGHT: In this work, we generate a low-dimensional representation from a short unit segment of audio, and couple this fingerprint with a fast maximum inner-product search.
- 607, TITLE: Self-Training and Pre-Training are Complementary for Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9414641>  
AUTHORS: Q. Xu et al.  
HIGHLIGHT: In this paper, we show that pseudo-labeling and pre-training with wav2vec 2.0 are complementary in a variety of labeled data setups.
- 608, TITLE: Unsupervised Discriminative Learning of Sounds for Audio Event Classification  
<https://doi.org/10.1109/ICASSP39728.2021.9413482>  
AUTHORS: S. Hornauer, K. Li, S. X. Yu, S. Ghaffarzadegan and L. Ren  
HIGHLIGHT: On several audio event classification benchmarks, we show a fast and effective alternative that pre-trains the model unsupervised, only on audio data and yet delivers on-par performance with ImageNet pre-training.
- 609, TITLE: Similarity Analysis of Self-Supervised Speech Representations  
<https://doi.org/10.1109/ICASSP39728.2021.9414321>  
AUTHORS: Y. -A. Chung, Y. Belinkov and J. Glass  
HIGHLIGHT: In this work, we aim to provide a comparative study of some of the most representative self-supervised algorithms.
- 610, TITLE: Joint Masked CPC And CTC Training For ASR  
<https://doi.org/10.1109/ICASSP39728.2021.9414227>  
AUTHORS: C. Talnikar, T. Likhomanenko, R. Collobert and G. Synnaeve  
HIGHLIGHT: In this paper we demonstrate a single-stage training of ASR models that can utilize both unlabeled and labeled data.
- 611, TITLE: A Comparison of Discrete Latent Variable Models for Speech Representation Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9413680>

AUTHORS: H. Zhou, A. Baevski and M. Auli  
HIGHLIGHT: This paper presents a comparison of two different approaches which are broadly based on predicting future time-steps or auto-encoding the input signal.

612, TITLE: Federated Learning from Big Data Over Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414903>  
AUTHORS: Y. Sarcheshmehpour, M. Leinonen and A. Jung  
HIGHLIGHT: This paper formulates and studies a novel algorithm for federated learning from large collections of local datasets.

613, TITLE: Efficient Client Contribution Evaluation for Horizontal Federated Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9413377>  
AUTHORS: J. Zhao, X. Zhu, J. Wang and J. Xiao  
HIGHLIGHT: In this paper an efficient method is proposed to evaluate the contributions of federated participants.

614, TITLE: A Quantitative Metric for Privacy Leakage in Federated Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9413539>  
AUTHORS: Y. Liu, X. Zhu, J. Wang and J. Xiao  
HIGHLIGHT: In this paper, we propose a novel method to approximate the mutual information between the high-dimensional gradients and batched input data.

615, TITLE: DP-SIGNSGD: When Efficiency Meets Privacy and Robustness  
<https://doi.org/10.1109/ICASSP39728.2021.9414538>  
AUTHORS: L. Lyu  
HIGHLIGHT: In this paper, we bridge this gap by presenting an improved method called DP-SIGNSGD, which can meet all the aforementioned properties.

616, TITLE: Federated Algorithm with Bayesian Approach: Omni-Fedge  
<https://doi.org/10.1109/ICASSP39728.2021.9413571>  
AUTHORS: S. A. Kesanapalli and B. N. Bharath  
HIGHLIGHT: In this paper, we consider the problem of Federated Learning (FL) under non-i.i.d data setting.

617, TITLE: Training Speech Recognition Models with Federated Learning: A Quality/Cost Framework  
<https://doi.org/10.1109/ICASSP39728.2021.9413397>  
AUTHORS: D. Guliani, F. Beaufays and G. Motta  
HIGHLIGHT: We propose a framework by which the degree of non-IID-ness can be varied, consequently illustrating a trade-off between model quality and the computational cost of federated training, which we capture through a novel metric.

618, TITLE: Cross-Silo Federated Training in the Cloud with Diversity Scaling and Semi-Supervised Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9413428>  
AUTHORS: K. Nandury, A. Mohan and F. Weber  
HIGHLIGHT: We propose a novel aggregation algorithm that takes update diversity into account and significantly outperforms Federated Averaging (FedAvg).

619, TITLE: Gradual Federated Learning Using Simulated Annealing  
<https://doi.org/10.1109/ICASSP39728.2021.9414455>  
AUTHORS: L. T. Nguyen and B. Shim  
HIGHLIGHT: In this work, we put forth a new update strategy based on the simulated annealing (SA) algorithm, in which the user devices choose their training parameters between the global evaluation model and their local models probabilistically.

620, TITLE: Optimal Importance Sampling for Federated Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9413655>  
AUTHORS: E. Rizk, S. Vlaski and A. H. Sayed  
HIGHLIGHT: We derive optimal importance sampling strategies for both agent and data selection and show that under convexity and Lipschitz assumptions, non-uniform sampling without replacement improves the performance of the original FedAvg algorithm.

621, TITLE: Multi-Tier Federated Learning for Vertically Partitioned Data  
<https://doi.org/10.1109/ICASSP39728.2021.9415026>  
AUTHORS: A. Das and S. Patterson

**HIGHLIGHT:** We propose Tiered Decentralized Coordinate Descent (TDCD), a communication-efficient decentralized training algorithm for such two-tiered networks.

622, **TITLE:** Energy Minimization for Federated Learning with IRS-Assisted Over-the-Air Computation  
<https://doi.org/10.1109/ICASSP39728.2021.9414785>  
**AUTHORS:** Y. Hu et al.  
**HIGHLIGHT:** This paper investigates the deployment of federated learning (FL) over an over-the-air computation (AirComp) and intelligent reflecting surface (IRS) based wireless network.

623, **TITLE:** Adaptive Quantization of Model Updates for Communication-Efficient Federated Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9413697>  
**AUTHORS:** D. Jhunjhunwala, A. Gadhikar, G. Joshi and Y. C. Eldar  
**HIGHLIGHT:** In this work, we propose an adaptive quantization strategy called AdaQuantFL that aims to achieve communication efficiency as well as a low error floor by changing the number of quantization levels during the course of training.

624, **TITLE:** HebbNet: A Simplified Hebbian Learning Framework to do Biologically Plausible Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9414241>  
**AUTHORS:** M. Gupta, A. Ambikapathi and S. Ramasamy  
**HIGHLIGHT:** In this work, we introduce a new Hebbian learning based neural network, called HebbNet.

625, **TITLE:** t-k-means: A ROBUST AND STABLE k-means VARIANT  
<https://doi.org/10.1109/ICASSP39728.2021.9414687>  
**AUTHORS:** Y. Li, Y. Zhang, Q. Tang, W. Huang, Y. Jiang and S. -T. Xia  
**HIGHLIGHT:** In this paper, we propose a robust and stable k-means variant, dubbed the t-k-means, as well as its fast version to alleviate those problems.

626, **TITLE:** Adaptive Feature Weight Learning For Robust Clustering Problem with Sparse Constraint  
<https://doi.org/10.1109/ICASSP39728.2021.9413845>  
**AUTHORS:** F. Nie, W. Chang, X. Li, J. Xu and G. Li  
**HIGHLIGHT:** To deal with this situation, we propose a novel weight learning mechanism to adaptively reweight each feature in the data.

627, **TITLE:** Assisted Learning: Cooperative AI with Autonomy  
<https://doi.org/10.1109/ICASSP39728.2021.9414618>  
**AUTHORS:** J. Zhou, X. Xian, N. Li and J. Ding  
**HIGHLIGHT:** In this paper, we propose a method named ASCII for an agent to improve its classification performance through assistance from other agents, without sharing proprietary data and model information.

628, **TITLE:** Geom-Spider-EM: Faster Variance Reduced Stochastic Expectation Maximization for Nonconvex Finite-Sum Optimization  
<https://doi.org/10.1109/ICASSP39728.2021.9414271>  
**AUTHORS:** G. Fort, E. Moulines and H. -T. Wai  
**HIGHLIGHT:** In this paper, we propose an extension of the Stochastic Path-Integrated Differential Estimator EM (SPIDER-EM) and derive complexity bounds for this novel algorithm, designed to solve smooth nonconvex finite-sum optimization problems.

629, **TITLE:** Learning a Tree of Neural Nets  
<https://doi.org/10.1109/ICASSP39728.2021.9413718>  
**AUTHORS:** A. Zharmagambetov and M. ?. Carreira-Perpi??n  
**HIGHLIGHT:** In particular, we show that, rather than improving a ResNet by making it deeper, it is better to construct a tree of small ResNets.

630, **TITLE:** Corrupted Contextual Bandits: Online Learning with Corrupted Context  
<https://doi.org/10.1109/ICASSP39728.2021.9414300>  
**AUTHORS:** D. Bouneffouf  
**HIGHLIGHT:** In order to address the corrupted-context setting, we propose to combine the standard contextual bandit approach with a classical multi-armed bandit mechanism.

631, **TITLE:** Training a Bank of Wiener Models with a Novel Quadratic Mutual Information Cost Function  
<https://doi.org/10.1109/ICASSP39728.2021.9415011>  
**AUTHORS:** B. Hu and J. C. Principe



**HIGHLIGHT:** This paper presents a novel training methodology to adapt parameters of a bank of Wiener models (BWMs), i.e., a bank of linear filters followed by a static memoryless nonlinearity, using full pdf information of the projected outputs and the desired signal.

632, **TITLE:** Information and Regularization in Restricted Boltzmann Machines

<https://doi.org/10.1109/ICASSP39728.2021.9414497>

**AUTHORS:** M. Vera, L. R. Vega and P. Piantanida

**HIGHLIGHT:** In this work, we study mutual information in Restricted Boltzmann Machines (RBM) and its relationship with the different regularization techniques.

633, **TITLE:** Deep Deterministic Information Bottleneck with Matrix-Based Entropy Functional

<https://doi.org/10.1109/ICASSP39728.2021.9414151>

**AUTHORS:** X. Yu, S. Yu and J. C. Principe

**HIGHLIGHT:** We introduce the matrix-based Rényi's  $\alpha$ -order entropy functional to parameterize Tishby et al. information bottleneck (IB) principle [1] with a neural network.

634, **TITLE:** Transitive Transfer Sparse Coding for Distant Domain

<https://doi.org/10.1109/ICASSP39728.2021.9415021>

**AUTHORS:** L. Feng, F. Qian, X. He, Y. Fan, H. Cai and G. Hu

**HIGHLIGHT:** In this paper, we propose a novel framework called transitive transfer sparse coding (TTSC) to solve the two distant domains transfer learning problem.

635, **TITLE:** Fast Local Representation Learning with Adaptive Anchor Graph

<https://doi.org/10.1109/ICASSP39728.2021.9414630>

**AUTHORS:** C. Zhang, F. Nie, Z. Wang, R. Wang and X. Li

**HIGHLIGHT:** In order to solve this issue, we propose a fast local representation learning with adaptive anchor graph to learn local structure information through similarity matrix in anchor-based graph.

636, **TITLE:** Learning On Heterogeneous Graphs Using High-Order Relations

<https://doi.org/10.1109/ICASSP39728.2021.9413417>

**AUTHORS:** S. H. Lee, F. Ji and W. P. Tay

**HIGHLIGHT:** In this paper, we propose an approach for learning on heterogeneous graphs without using meta-paths.

637, **TITLE:** Incomplete Multi-View Subspace Clustering with Low-Rank Tensor

<https://doi.org/10.1109/ICASSP39728.2021.9414688>

**AUTHORS:** J. Liu, S. Teng, W. Zhang, X. Fang, L. Fei and Z. Zhang

**HIGHLIGHT:** To address these issues, we propose a novel Incomplete Multi-view Subspace Clustering with Low-rank Tensor (IMSCLT) method, which could be the first tensor-based incomplete multi-view clustering method to the best of our knowledge.

638, **TITLE:** Channel-Wise Mix-Fusion Deep Neural Networks for Zero-Shot Learning

<https://doi.org/10.1109/ICASSP39728.2021.9413569>

**AUTHORS:** G. Wang, N. Guau, H. Ye, X. Yi, H. Cheng and J. Zhu

**HIGHLIGHT:** This paper proposes a channel-wise mix-fusion ZSL model (CMFZ) to contextualize the ZSL classifier's discriminative information by incorporating much richer visual semantic information from both objects and their semantic surrounding environments.

639, **TITLE:** Online Unsupervised Learning Using Ensemble Gaussian Processes with Random Features

<https://doi.org/10.1109/ICASSP39728.2021.9414960>

**AUTHORS:** G. V. Karanikolas, Q. Lu and G. B. Giannakis

**HIGHLIGHT:** This work develops an efficient online approach based on random features by replacing spatial with spectral subsampling.

640, **TITLE:** Dimension Selected Subspace Clustering

<https://doi.org/10.1109/ICASSP39728.2021.9413849>

**AUTHORS:** S. Li, Y. Luo, J. Chambers and W. Wang

**HIGHLIGHT:** In this paper, a new dimension sketching algorithm is proposed, which aims to select informative dimensions that have significant effects on the clustering results.

641, **TITLE:** Deep Ensemble Siamese Network For Incremental Signal Classification

<https://doi.org/10.1109/ICASSP39728.2021.9414902>

AUTHORS: C. Yang and S. Yang  
HIGHLIGHT: In this paper, a new Deep Ensemble Siamese Network (DESN) is constructed for unknown category detection and incremental accumulation of signals from the detected category.

642, TITLE: Non-Recursive Graph Convolutional Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414106>  
AUTHORS: H. Chen, Z. Deng, Y. Xu and Z. Li  
HIGHLIGHT: Therefore, in this paper, we propose a novel architecture named Non-Recursive Graph Convolutional Network (NRGCN) to improve both the training efficiency and the learning performance of GCNs in the context of node classification.

643, TITLE: Ego-Based Entropy Measures for Structural Representations on Graphs  
<https://doi.org/10.1109/ICASSP39728.2021.9414279>  
AUTHORS: G. Dasoulas, G. Nikolentzos, K. Seaman, A. Virmaux and M. Vazirgiannis  
HIGHLIGHT: In this paper, we propose VNEstruct, a simple approach, based on entropy measures of the neighborhood's topology, for generating low-dimensional structural representations, that is time- efficient and robust to graph perturbations.

644, TITLE: Symmetric Sub-graph Spatio-Temporal Graph Convolution and its application in Complex Activity Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9413833>  
AUTHORS: P. Das and A. Ortega  
HIGHLIGHT: In this paper, we analyze hand skeleton-based complex activities by modeling dynamic hand skeletons through a spatiotemporal graph convolutional neural network (ST-GCN).

645, TITLE: Progressive Spatio-Temporal Graph Convolutional Network for Skeleton-Based Human Action Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9413860>  
AUTHORS: N. Heidari and A. Iosifidis  
HIGHLIGHT: In this paper, we propose a method to automatically find a compact and problem-specific topology for spatio-temporal graph convolutional networks in a progressive manner.

646, TITLE: Sparse-Coded Dynamic Mode Decomposition on Graph for Prediction of River Water Level Distribution  
<https://doi.org/10.1109/ICASSP39728.2021.9414533>  
AUTHORS: Y. ARAI, S. MURAMATSU, H. YASUDA, K. HAYASAKA and Y. OTAKE  
HIGHLIGHT: This work proposes a method for estimating dynamics on graph by using dynamic mode decomposition (DMD) and sparse approximation with graph filter banks (GFBs).

647, TITLE: Graph Frequency Analysis of COVID-19 Incidence to Identify County-Level Contagion Patterns in the United States  
<https://doi.org/10.1109/ICASSP39728.2021.9414854>  
AUTHORS: Y. Li and G. Mateos  
HIGHLIGHT: We thus conduct a graph frequency analysis to investigate the spread patterns of COVID-19 in different US counties.

648, TITLE: Generalized Polytopic Matrix Factorization  
<https://doi.org/10.1109/ICASSP39728.2021.9413709>  
AUTHORS: G. Tatli and A. T. Erdogan  
HIGHLIGHT: This article aims to characterize all eligible polytopic sets that are suitable for the PMF framework.

649, TITLE: Exact Linear Convergence Rate Analysis for Low-Rank Symmetric Matrix Completion via Gradient Descent  
<https://doi.org/10.1109/ICASSP39728.2021.9413419>  
AUTHORS: T. Vu and R. Raich  
HIGHLIGHT: This paper performs a local analysis of the exact linear convergence rate of gradient descent for factorization-based symmetric matrix completion.

650, TITLE: Structured Support Exploration for Multilayer Sparse Matrix Factorization  
<https://doi.org/10.1109/ICASSP39728.2021.9414238>  
AUTHORS: Q. -T. Le and R. Gribonval  
HIGHLIGHT: In this paper, we address two problems with the application of proximal algorithms to sparse matrix factorization.

651, TITLE: Optimal Selection of Matrix Shape and Decomposition Scheme for Neural Network Compression

<https://doi.org/10.1109/ICASSP39728.2021.9414224>

AUTHORS: Y. Idelbayev and M. ?. Carreira-Perpi??n

HIGHLIGHT: We formulate this problem as a mixed-integer optimization over the weights, ranks, and decompositions schemes; and we provide an efficient alternating optimization algorithm involving two simple steps: a step over the weights of the neural network (solved by SGD), and a step over the ranks and decomposition schemes (solved by an SVD).

652, TITLE: Sparse Graph Based Sketching for Fast Numerical Linear Algebra

<https://doi.org/10.1109/ICASSP39728.2021.9414030>

AUTHORS: D. Hu, S. Ubaru, A. Gittens, K. L. Clarkson, L. Horesh and V. Kalantzis

HIGHLIGHT: In this paper, we study sketching matrices that are obtained from bipartite graphs that are sparse, i.e., have left degree  $s$  that is small.

653, TITLE: Cold Start Revisited: A Deep Hybrid Recommender with Cold-Warm Item Harmonization

<https://doi.org/10.1109/ICASSP39728.2021.9413384>

AUTHORS: O. Barkan, R. Hirsch, O. Katz, A. Caciularu, Y. Weill and N. Koenigstein

HIGHLIGHT: In this paper, we show that a straightforward application of parametric models may lead to discrepancies between the cold and warm items' distributions in the CF space.

654, TITLE: On a Guided Nonnegative Matrix Factorization

<https://doi.org/10.1109/ICASSP39728.2021.9413656>

AUTHORS: J. Vendrow, J. Haddock, E. Rebrova and D. Needell

HIGHLIGHT: For this reason, we propose an approach based upon the nonnegative matrix factorization (NMF) model, deemed Guided NMF, that incorporates user-designed seed word supervision.

655, TITLE: Nonnegative Unimodal Matrix Factorization

<https://doi.org/10.1109/ICASSP39728.2021.9414631>

AUTHORS: A. M. S. Ang, N. Gillis, A. Vandaele and H. D. Sterck

HIGHLIGHT: We introduce a new Nonnegative Matrix Factorization (NMF) model called Nonnegative Unimodal Matrix Factorization (NuMF), which adds on top of NMF the unimodal condition on the columns of the basis matrix.

656, TITLE: Kernel Orthogonal Nonnegative Matrix Factorization: Application to Multispectral Document Image Decomposition

<https://doi.org/10.1109/ICASSP39728.2021.9413964>

AUTHORS: A. Rahiche and M. Cheriet

HIGHLIGHT: In this work, we propose a new kernel orthogonal NMF method that does not suffer from the pre-image issue.

657, TITLE: Random Projection Streams for (Weighted) Nonnegative Matrix Factorization

<https://doi.org/10.1109/ICASSP39728.2021.9413496>

AUTHORS: F. Yahaya, M. Puigt, G. Delmaire and G. Roussel

HIGHLIGHT: In this paper, we thus investigate an alternative framework to structured random projections?named random projection streams (RPS)?which (i) are based on classical random compression strategies only?and are thus data-independent?and (ii) can benefit from the above fast techniques.

658, TITLE: Multivariate Non-Negative Matrix Factorization with Application to Energy Disaggregation

<https://doi.org/10.1109/ICASSP39728.2021.9414202>

AUTHORS: P. A. Schirmer and I. Mporas

HIGHLIGHT: In the proposed architecture the baseline Non-negative Matrix Factorization approach is expanded utilizing multi-variate signals.

659, TITLE: Continuous-Time Self-Attention in Neural Differential Equation

<https://doi.org/10.1109/ICASSP39728.2021.9414104>

AUTHORS: J. -T. Chien and Y. -H. Chen

HIGHLIGHT: This study presents a new continuous-time attention to improve sequential learning where the region of interest in continuous-time state trajectory over observed as well as missing samples is sufficiently attended.

660, TITLE: Blind Deinterleaving of Signals in Time Series with Self-Attention Based Soft Min-Cost Flow Learning

<https://doi.org/10.1109/ICASSP39728.2021.9415025>

AUTHORS: O. Can, Y. Z. G?rb?z, B. Yildirim and A. A. Alatan

HIGHLIGHT: We propose an end-to-end learning approach to address deinterleaving of patterns in time series, in particular, radar signals.

- 661, TITLE: Attention on Attention Sparse Dense Convolutional Network for Financial Signal Processing  
<https://doi.org/10.1109/ICASSP39728.2021.9414397>  
AUTHORS: T. Zhu, J. Li, X. Liu, Y. Jiang and S. -T. Xia  
HIGHLIGHT: For a better solution, we propose a novel Attention on Attention Sparse Dense Convolutional Network (AoA-SDCN), which strengthens time decay characteristics by adding dense connections at close points.
- 662, TITLE: Stock Movement Prediction and Portfolio Management via Multimodal Learning with Transformer  
<https://doi.org/10.1109/ICASSP39728.2021.9414893>  
AUTHORS: D. Daiya and C. Lin  
HIGHLIGHT: This paper introduces a novel high performing multimodal deep learning architecture(Trans-DiCE) for stock movement prediction utilizing financial indicators and news data.
- 663, TITLE: A Quaternion-Valued Variational Autoencoder  
<https://doi.org/10.1109/ICASSP39728.2021.9413859>  
AUTHORS: E. Grassucci, D. Comminiello and A. Uncini  
HIGHLIGHT: In this paper, we propose a novel VAE defined in the quaternion domain, which exploits the properties of quaternion algebra to improve performance while significantly reducing the number of parameters required by the network.
- 664, TITLE: Learning a Sparse Generative Non-Parametric Supervised Autoencoder  
<https://doi.org/10.1109/ICASSP39728.2021.9414410>  
AUTHORS: M. Barlaud and F. Guyard  
HIGHLIGHT: This paper concerns the supervised generative non parametric autoencoder.
- 665, TITLE: DAG-GAN: Causal Structure Learning with Generative Adversarial Nets  
<https://doi.org/10.1109/ICASSP39728.2021.9414770>  
AUTHORS: Y. Gao, L. Shen and S. -T. Xia  
HIGHLIGHT: In this paper, we consider DAG structure learning from the perspective of distributional optimization and design an adversarial framework named DAG-GAN to detect the DAG structure from data.
- 666, TITLE: Relaxed Wasserstein with Applications to GANs  
<https://doi.org/10.1109/ICASSP39728.2021.9414454>  
AUTHORS: X. Guo, J. Hong, T. Lin and N. Yang  
HIGHLIGHT: In this paper, we propose a new class of Relaxed Wasserstein (RW) distances by generalizing Wasserstein-1 distance with Bregman cost functions.
- 667, TITLE: Environment-Independent Wi-Fi Human Activity Recognition with Adversarial Network  
<https://doi.org/10.1109/ICASSP39728.2021.9413590>  
AUTHORS: Z. Wang, S. Chen, W. Yang and Y. Xu  
HIGHLIGHT: To address this issue, we in this paper present WiHARAN, a Wi-Fi-based activity recognition system that can learn environment-independent features from Channel State Information (CSI) traces.
- 668, TITLE: A Robust to Noise Adversarial Recurrent Model for Non-Intrusive Load Monitoring  
<https://doi.org/10.1109/ICASSP39728.2021.9413663>  
AUTHORS: M. Kaselimi, A. Voulodimos, N. Doulamis, A. Doulamis and E. Protopapadakis  
HIGHLIGHT: In this paper, we propose EnerGAN++, an adversarially trained model for robust energy disaggregation.
- 669, TITLE: Enhancing Data-Free Adversarial Distillation with Activation Regularization and Virtual Interpolation  
<https://doi.org/10.1109/ICASSP39728.2021.9413483>  
AUTHORS: X. Qu, J. Wang and J. Xiao  
HIGHLIGHT: We add an activation regularizer and a virtual interpolation method to improve the data generation efficiency.
- 670, TITLE: Sequential Adversarial Anomaly Detection with Deep Fourier Kernel  
<https://doi.org/10.1109/ICASSP39728.2021.9414197>  
AUTHORS: S. Zhu, H. S. Yuchi, M. Zhang and Y. Xie  
HIGHLIGHT: We present a novel adversarial detector for the anomalous sequence when there are only one-class training samples.
- 671, TITLE: Incorporate Maximum Mean Discrepancy in Recurrent Latent Space for Sequential Generative Model  
<https://doi.org/10.1109/ICASSP39728.2021.9414580>

AUTHORS: Y. Zhang, Y. Wang and Y. Dong  
HIGHLIGHT: In this paper, we incorporate maximum mean discrepancy in the recurrent structure for distribution regularization.

672, TITLE: FMA-ETA: Estimating Travel Time Entirely Based on FFN with Attention  
<https://doi.org/10.1109/ICASSP39728.2021.9414054>  
AUTHORS: Y. Sun et al.  
HIGHLIGHT: To solve this problem, we propose a novel, brief and effective framework mainly based on feed-forward network (FFN) for ETA, FFN with Multifactor Attention (FMA-ETA).

673, TITLE: A Unified Approach to Translate Classical Bandit Algorithms to Structured Bandits  
<https://doi.org/10.1109/ICASSP39728.2021.9413628>  
AUTHORS: S. Gupta, S. Chaudhari, S. Mukherjee, G. Joshi and O. Yagan  
HIGHLIGHT: We propose a novel approach to gradually estimate the hidden  $\theta^*$  and use the estimate together with the mean reward functions to substantially reduce exploration of sub-optimal arms.

674, TITLE: Near-Optimal Algorithms for Piecewise-Stationary Cascading Bandits  
<https://doi.org/10.1109/ICASSP39728.2021.9414506>  
AUTHORS: L. Wang, H. Zhou, B. Li, L. R. Varshney and Z. Zhao  
HIGHLIGHT: Considering piecewise-stationary environments, two efficient algorithms, GLRT-CascadeUCB and GLRT-CascadeKL-UCB, are developed.

675, TITLE: Optimum Feature Ordering for Dynamic Instance-Wise Joint Feature Selection and Classification  
<https://doi.org/10.1109/ICASSP39728.2021.9414669>  
AUTHORS: Y. W. Liyanage and D. Zois  
HIGHLIGHT: Based on these properties, we propose a fast algorithm and demonstrate its effectiveness compared to the state-of-the-art using 4 publicly available datasets.

676, TITLE: POLA: Online Time Series Prediction by Adaptive Learning Rates  
<https://doi.org/10.1109/ICASSP39728.2021.9414906>  
AUTHORS: W. Zhang  
HIGHLIGHT: We propose POLA (Predicting Online by Learning rate Adaptation) to automatically regulate the learning rate of recurrent neural network models to adapt to changing time series patterns across time.

677, TITLE: Singer Identification Using Deep Timbre Feature Learning with KNN-NET  
<https://doi.org/10.1109/ICASSP39728.2021.9413774>  
AUTHORS: X. Zhang, J. Qian, Y. Yu, Y. Sun and W. Li  
HIGHLIGHT: In this paper, we study the issue of automatic singer identification (SID) in popular music recordings, which aims to recognize who sang a given piece of song.

678, TITLE: Implicit HRTF Modeling Using Temporal Convolutional Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414750>  
AUTHORS: I. D. Gebru et al.  
HIGHLIGHT: In this work, we present a data-driven approach to learn HRTFs implicitly with a neural network that achieves state of the art results compared to traditional approaches but relies on a much simpler data capture that can be performed in arbitrary, non-anechoic rooms.

679, TITLE: Improving the Classification of Rare Chords With Unlabeled Data  
<https://doi.org/10.1109/ICASSP39728.2021.9413701>  
AUTHORS: M. Bortolozzo, R. Schramm and C. R. Jung  
HIGHLIGHT: In this work, we explore techniques to improve performance for rare classes in the task of Automatic Chord Recognition (ACR).

680, TITLE: Loopnet: Musical Loop Synthesis Conditioned on Intuitive Musical Parameters  
<https://doi.org/10.1109/ICASSP39728.2021.9415047>  
AUTHORS: P. Chandna, A. Ramires, X. Serra and E. Gómez  
HIGHLIGHT: Taking such criteria into account, we present LoopNet, a feed-forward generative model for creating loops conditioned on intuitive parameters.

681, TITLE: Micaugment: One-Shot Microphone Style Transfer

<https://doi.org/10.1109/ICASSP39728.2021.9413994>

AUTHORS: Z. Borsos, Y. Li, B. Gfeller and M. Tagliasacchi

HIGHLIGHT: In this work, we propose a method to perform one-shot microphone style transfer.

682, TITLE: Wasserstein Barycenter Transport for Acoustic Adaptation

<https://doi.org/10.1109/ICASSP39728.2021.9414199>

AUTHORS: E. F. Montesuma and F. -M. Ngol? Mboula

HIGHLIGHT: This paper proposes a multi-source domain adaptation algorithm called Wasserstein Barycenter Transport, which transports the source domains to a target domain by creating an intermediate domain using the Wasserstein barycenter.

683, TITLE: Efficient Adversarial Audio Synthesis VIA Progressive Upsampling

<https://doi.org/10.1109/ICASSP39728.2021.9413954>

AUTHORS: Y. Cho, M. Chang, S. Lee, H. Lee, G. J. Kim and J. Choo

HIGHLIGHT: This paper proposes a novel generative model called PUGAN, which progressively synthesizes high-quality audio in a raw waveform.

684, TITLE: Multi-Channel Speech Enhancement Using Graph Neural Networks

<https://doi.org/10.1109/ICASSP39728.2021.9413955>

AUTHORS: P. Tzirakis, A. Kumar and J. Donley

HIGHLIGHT: In this paper, we introduce a different research direction by viewing each audio channel as a node lying in a non-Euclidean space and, specifically, a graph.

685, TITLE: Multi-Decoder Dprnn: Source Separation for Variable Number of Speakers

<https://doi.org/10.1109/ICASSP39728.2021.9414205>

AUTHORS: J. Zhu, R. A. Yeh and M. Hasegawa-Johnson

HIGHLIGHT: We propose an end-to-end trainable approach to single-channel speech separation with unknown number of speakers.

686, TITLE: Data-Efficient Framework for Real-World Multiple Sound Source 2d Localization

<https://doi.org/10.1109/ICASSP39728.2021.9413695>

AUTHORS: G. Le Moing et al.

HIGHLIGHT: We propose to use adversarial learning methods to close the gap between synthetic and real do-mains.

687, TITLE: Fusing Information Streams in End-to-End Audio-Visual Speech Recognition

<https://doi.org/10.1109/ICASSP39728.2021.9414553>

AUTHORS: W. Yu, S. Zeiler and D. Kolossa

HIGHLIGHT: We propose a new fusion strategy, incorporating reliability information in a decision fusion net that considers the temporal effects of the attention mechanism.

688, TITLE: Cooperative Scenarios for Multi-Agent Reinforcement Learning in Wireless Edge Caching

<https://doi.org/10.1109/ICASSP39728.2021.9414319>

AUTHORS: N. Garg and T. Ratnarajah

HIGHLIGHT: In this paper, we investigate multi-agent reinforcement learning (MARL), and identify four scenarios for cooperation.

689, TITLE: Robust Deep Reinforcement Learning for Underwater Navigation with Unknown Disturbances

<https://doi.org/10.1109/ICASSP39728.2021.9414937>

AUTHORS: J. Parras and S. Zazo

HIGHLIGHT: We propose a robust Proximal Policy Optimization agent and train it using simulations of an underwater medium: this agent shows an excellent performance when facing unknown disturbances, being able to approach the performance of the optimal agent which had an exact knowledge of the underwater disturbance.

690, TITLE: Online Hyper-Parameter Tuning for the Contextual Bandit

<https://doi.org/10.1109/ICASSP39728.2021.9413526>

AUTHORS: D. Bouneffouf and E. Claeys

HIGHLIGHT: We have presented here two algorithms that uses a bandit to find the optimal exploration of the contextual bandit algorithm, which we hope is the first step toward the automation of the multi-armed bandit algorithm.

691, TITLE: Double-Linear Thompson Sampling for Context-Attentive Bandits

<https://doi.org/10.1109/ICASSP39728.2021.9413882>

**AUTHORS:** D. Bouneffouf, R. Feraud, S. Upadhyay, Y. Khazaeni and I. Rish  
**HIGHLIGHT:** In this paper, we analyze and extend an online learning frame-work known as Context-Attentive Bandit, motivated by various practical applications, from medical diagnosis to dialog systems, where due to observation costs only a small subset of a potentially large number of context variables can be observed at each iteration; however, the agent has a freedom to choose which variables to observe.

692, **TITLE:** On the Marginal Benefit of Active Learning: Does Self-Supervision Eat its Cake?  
<https://doi.org/10.1109/ICASSP39728.2021.9414665>  
**AUTHORS:** Y. -C. Chan, M. Li and S. Oymak  
**HIGHLIGHT:** To this aim, this paper provides a novel algorithmic framework integrating self-supervised pretraining, active learning, and consistency-regularized self-training.

693, **TITLE:** Robust Maml: Prioritization Task Buffer with Adaptive Learning Process for Model-Agnostic Meta-Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9413446>  
**AUTHORS:** T. Nguyen, T. Luu, T. Pham, S. Rakhimkul and C. D. Yoo  
**HIGHLIGHT:** This paper proposes a more robust MAML based on an adaptive learning scheme and a prioritization task buffer (PTB) referred to as Robust MAML (RMAML) for improving scalability of training process and alleviating the problem of distribution mismatch.

694, **TITLE:** Introducing Deep Reinforcement Learning to Nlu Ranking Tasks  
<https://doi.org/10.1109/ICASSP39728.2021.9414475>  
**AUTHORS:** G. Yu, E. Barut and C. Su  
**HIGHLIGHT:** In this paper, we address these issues with a deep reinforcement learning approach that ranks suggestions from multiple experts in an online fashion.

695, **TITLE:** Temporal Link Prediction Via Reinforcement Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9413413>  
**AUTHORS:** Y. Tao, Y. Li and Z. Wu  
**HIGHLIGHT:** To deal with these challenges, we present a novel reinforcement learning approach with an update mechanism to integrate temporal information.

696, **TITLE:** A Deep Reinforcement Learning Approach To Audio-Based Navigation In A Multi-Speaker Environment  
<https://doi.org/10.1109/ICASSP39728.2021.9415013>  
**AUTHORS:** P. Giannakopoulos, A. Pikrakis and Y. Cotronis  
**HIGHLIGHT:** In this work we use deep reinforcement learning to create an autonomous agent that can navigate in a two-dimensional space using only raw auditory sensory information from the environment, a problem that has received very little attention in the reinforcement learning literature.

697, **TITLE:** Global-Localized Agent Graph Convolution for Multi-Agent Reinforcement Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9414993>  
**AUTHORS:** Y. Liu, Y. Dou, S. Shen and P. Qiao  
**HIGHLIGHT:** In this paper, we model the global and localized cooperation among agents by global and localized agent graphs and propose a novel graph convolutional reinforcement learning mechanism based on these two graphs which allows each agent to communicate with neighbors and all a-gents to cooperate at the high level.

698, **TITLE:** Gaussian Process Temporal-Difference Learning with Scalability and Worst-Case Performance Guarantees  
<https://doi.org/10.1109/ICASSP39728.2021.9414667>  
**AUTHORS:** Q. Lu and G. B. Giannakis  
**HIGHLIGHT:** The present paper revisits policy evaluation via temporal-difference (TD) learning from the Gaussian process (GP) perspective.

699, **TITLE:** Self-Inference Of Others? Policies For Homogeneous Agents In Cooperative Multi-Agent Reinforcement Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9414712>  
**AUTHORS:** Q. Lin and Q. Ling  
**HIGHLIGHT:** To address this issue, we propose to let each agent infer the others' policies with its own model, given that the agents are homogeneous.

700, **TITLE:** Semi-Supervised Batch Active Learning Via Bilevel Optimization  
<https://doi.org/10.1109/ICASSP39728.2021.9414206>  
**AUTHORS:** Z. Borsos, M. Tagliasacchi and A. Krause



**HIGHLIGHT:** In this work, we propose a novel batch acquisition strategy for active learning in the setting where the model training is performed in a semi-supervised manner.

701, **TITLE:** Kernel-Based Lifelong Policy Gradient Reinforcement Learning

<https://doi.org/10.1109/ICASSP39728.2021.9414511>

**AUTHORS:** R. Mowakeaa, S. -J. Kim and D. K. Emge

**HIGHLIGHT:** In this work, we propose a lifelong RL algorithm based on the kernel method to leverage nonlinear features of the data based on a popular union-of-subspace model.

702, **TITLE:** Policy Augmentation: An Exploration Strategy For Faster Convergence of Deep Reinforcement Learning Algorithms

<https://doi.org/10.1109/ICASSP39728.2021.9413545>

**AUTHORS:** A. Mahyari

**HIGHLIGHT:** In this paper, a revolutionary algorithm, called Policy Augmentation, is introduced.

703, **TITLE:** Graphcomm: A Graph Neural Network Based Method for Multi-Agent Reinforcement Learning

<https://doi.org/10.1109/ICASSP39728.2021.9413716>

**AUTHORS:** S. Shen et al.

**HIGHLIGHT:** In this work, we propose GraphComm, a method makes use of the relation-ships among agents for MARL communication.

704, **TITLE:** In Situ Calibration of Cross-Sensitive Sensors in Mobile Sensor Arrays Using Fast Informed Non-Negative Matrix Factorization

<https://doi.org/10.1109/ICASSP39728.2021.9414742>

**AUTHORS:** O. V. thanh, M. Puigt, F. Yahaya, G. Delmaire and G. Roussel

**HIGHLIGHT:** In this paper, we assume a set of mobile geolocalized sensor arrays observing an area over time.

705, **TITLE:** Multiphish: Multi-Modal Features Fusion Networks for Phishing Detection

<https://doi.org/10.1109/ICASSP39728.2021.9415016>

**AUTHORS:** L. Zhang, P. Zhang, L. Liu and J. Tan

**HIGHLIGHT:** In this paper, we propose a features fusion networks (MultiPhish) which is the first study on fusing multi-modal features with neural networks for the phishing detection task.

706, **TITLE:** Failure Prediction by Confidence Estimation of Uncertainty-Aware Dirichlet Networks

<https://doi.org/10.1109/ICASSP39728.2021.9414153>

**AUTHORS:** T. Tsiligkaridis

**HIGHLIGHT:** In this paper, it is first shown that uncertainty-aware deep Dirichlet neural networks provide an improved separation between the confidence of correct and incorrect predictions in the true class probability (TCP) metric.

707, **TITLE:** Two-Stage Framework for Seasonal Time Series Forecasting

<https://doi.org/10.1109/ICASSP39728.2021.9414118>

**AUTHORS:** Q. Xu, Q. Wen and L. Sun

**HIGHLIGHT:** In this paper, we propose a two-stage framework to forecast univariate seasonal time series.

708, **TITLE:** Recursive Input and State Estimation: a General Framework for Learning from Time Series With Missing Data

<https://doi.org/10.1109/ICASSP39728.2021.9414801>

**AUTHORS:** A. Garcia-Durán and R. West

**HIGHLIGHT:** In this paper, we introduce a single unifying framework, Recursive Input and State Estimation (RISE), for this general approach and reformulate existing models as specific instances of this framework.

709, **TITLE:** On the Performance-Complexity Tradeoff in Stochastic Greedy Weak Submodular Optimization

<https://doi.org/10.1109/ICASSP39728.2021.9413990>

**AUTHORS:** A. Hashemi, H. Vikalo and G. de Veciana

**HIGHLIGHT:** In this work, we study the tradeoff between performance and complexity when one resorts to random sampling strategies to reduce the query complexity of GREEDY.

710, **TITLE:** Semi-Supervised Time Series Classification by Temporal Relation Prediction

<https://doi.org/10.1109/ICASSP39728.2021.9413883>

**AUTHORS:** H. Fan, F. Zhang, R. Wang, X. Huang and Z. Li

**HIGHLIGHT:** In this work, we propose a simple and effective method of Semi-supervised Time series classification architecture (termed as SemiTime) by gaining from the structure of unlabeled data in a self-supervised manner.

711, TITLE: Continuous Cnn For Nonuniform Time Series

<https://doi.org/10.1109/ICASSP39728.2021.9414318>

AUTHORS: H. Shi et al.

**HIGHLIGHT:** In this paper, we propose the Continuous CNN (CCNN), which estimates the inherent continuous inputs by interpolation, and performs continuous convolution on the continuous input.

712, TITLE: Blend-Res2net: Blended Representation Space by Transformation of Residual Mapping with Restrained Learning for Time Series Classification

<https://doi.org/10.1109/ICASSP39728.2021.9414647>

AUTHORS: A. Ukil, A. J. Jara and L. Marin

**HIGHLIGHT:** We propose Blend-Res2Net that blends two different representation spaces:  $H1(x) = F(x) + \text{Trans}(x)$  and  $H2(x) = F(\text{Trans}(x)) + x$  with the intention of learning over richer representation by capturing the temporal as well as the spectral signatures ( $\text{Trans}(?)$  represents the transformation function).

713, TITLE: Spatiotemporal Attention for Multivariate Time Series Prediction and Interpretation

<https://doi.org/10.1109/ICASSP39728.2021.9413914>

AUTHORS: T. Gangopadhyay, S. Y. Tan, Z. Jiang, R. Meng and S. Sarkar

**HIGHLIGHT:** We propose a novel deep learning architecture, called spatiotemporal attention mechanism (STAM) for simultaneous learning of the most important time steps and variables.

714, TITLE: Tabular Transformers for Modeling Multivariate Time Series

<https://doi.org/10.1109/ICASSP39728.2021.9414142>

AUTHORS: I. Padhi et al.

**HIGHLIGHT:** Here we propose neural network models that represent tabular time series that can optionally leverage their hierarchical structure.

715, TITLE: Real-Time Synchronization in Neural Networks for Multivariate Time Series Anomaly Detection

<https://doi.org/10.1109/ICASSP39728.2021.9413847>

AUTHORS: A. Abdulaal and T. Lancewicki

**HIGHLIGHT:** We propose a mathematical formulation of neural network layers, which generate a synchronized representation from asynchronous multivariate input.

716, TITLE: Fast Graph Kernel with Optical Random Features

<https://doi.org/10.1109/ICASSP39728.2021.9413614>

AUTHORS: H. Ghanem, N. Keriven and N. Tremblay

**HIGHLIGHT:** In this paper, we propose to leverage kernel random features within the graphlet framework, and establish a theoretical link with a mean kernel metric.

717, TITLE: Fast Hierarchy Preserving Graph Embedding via Subspace Constraints

<https://doi.org/10.1109/ICASSP39728.2021.9414919>

AUTHORS: X. Chen, L. Du, M. Chen, Y. Wang, Q. Long and K. Xie

**HIGHLIGHT:** In this paper, we propose an inductive method, FastHGE, to learn node representations more efficiently and generalize to new nodes more easily.

718, TITLE: Graph Embedding using Multi-Layer Adjacent Point Merging Model

<https://doi.org/10.1109/ICASSP39728.2021.9413362>

AUTHORS: J. Huang and H. Kasai

**HIGHLIGHT:** To this end, we propose a novel graph embedding method using a multi-layer adjacent point merging model.

719, TITLE: Node Attribute Completion in Knowledge Graphs with Multi-Relational Propagation

<https://doi.org/10.1109/ICASSP39728.2021.9414016>

AUTHORS: E. Bayram, A. Garc?a-Dur?n and R. West

**HIGHLIGHT:** Our approach, denoted as MRAP, imputes the values of missing attributes by propagating information across the multi-relational structure of a knowledge graph.

720, TITLE: UserReg: A Simple but Strong Model for Rating Prediction

<https://doi.org/10.1109/ICASSP39728.2021.9413646>

AUTHORS: H. Zhang, I. Ganchev, N. S. Nikolov and M. Stevenson  
HIGHLIGHT: This paper proposes a simple linear model based on Matrix Factorization (MF), called UserReg, which regularizes users' latent representations with explicit feedback information for rating prediction.

721, TITLE: Toward Skills Dialog Orchestration with Online Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9414303>  
AUTHORS: D. Bouneffouf, R. Feraud, S. Upadhyay, M. Agarwal, Y. Khazaeni and I. Rish  
HIGHLIGHT: In this work, we study the task of online posterior dialog orchestration, where we define posterior orchestration as the task of selecting a subset of skills which most appropriately answer a user input using features extracted from both the user input and the individual skills.

722, TITLE: Adaptive Re-Balancing Network with Gate Mechanism for Long-Tailed Visual Question Answering  
<https://doi.org/10.1109/ICASSP39728.2021.9414074>  
AUTHORS: H. Chen, R. Liu, H. Fang and X. Zhang  
HIGHLIGHT: Therefore, we propose a unified Adaptive Re-balancing Network (ARN) to take care of classification in both head and tail classes, exhaustively improving performance for VQA.

723, TITLE: Co-Capsule Networks Based Knowledge Transfer for Cross-Domain Recommendation  
<https://doi.org/10.1109/ICASSP39728.2021.9413924>  
AUTHORS: H. Li, L. Yu, Y. Leng and Q. Du  
HIGHLIGHT: In this paper, we propose a CDR method with co-capsule networks based knowledge transfer to implement the recommendation for the cold-start users.

724, TITLE: Modurec: Recommender Systems with Feature and Time Modulation  
<https://doi.org/10.1109/ICASSP39728.2021.9413676>  
AUTHORS: J. Maroto, C. Vignac and P. Frossard  
HIGHLIGHT: To address these problems, we propose Modurec: an autoencoder-based method that combines all available information using the feature-wise modulation mechanism, which has demonstrated its effectiveness in several fields.

725, TITLE: Sig2Sig: Signal Translation Networks to Take the Remains of the Past  
<https://doi.org/10.1109/ICASSP39728.2021.9415084>  
AUTHORS: S. Kim, H. Lee, J. Han and J. -H. Kim  
HIGHLIGHT: In this paper, we propose a signal translation networks, Sig2Sig, that converts from the new sensor signals to old ones in order to reuse the past model, which was trained on plenty of old sensor signals.

726, TITLE: Solving a Class of Non-Convex Min-Max Games Using Adaptive Momentum Methods  
<https://doi.org/10.1109/ICASSP39728.2021.9414476>  
AUTHORS: B. Barazandeh, D. A. Tarzanagh and G. Michailidis  
HIGHLIGHT: In this paper, we propose an adaptive momentum min-max algorithm that generalizes adaptive momentum methods to the non-convex min-max regime.

727, TITLE: Minimizing Weighted Concave Impurity Partition Under Constraints  
<https://doi.org/10.1109/ICASSP39728.2021.9414887>  
AUTHORS: T. Nguyen and T. Nguyen  
HIGHLIGHT: In this paper, we formulate and solve a variant of the partition problem called the minimum weighted concave impurity partition under constraint (MIPUC).

728, TITLE: Constant Approximation Algorithm for Minimizing Concave Impurity  
<https://doi.org/10.1109/ICASSP39728.2021.9415005>  
AUTHORS: T. Nguyen, H. Le and T. Nguyen  
HIGHLIGHT: In this paper, we propose a linear time algorithm with bounded guarantee based on the maximum likelihood principle.

729, TITLE: Fusing Multitask Models by Recursive Least Squares  
<https://doi.org/10.1109/ICASSP39728.2021.9414440>  
AUTHORS: X. Li, L. Shan and W. Wang  
HIGHLIGHT: In this paper, we propose a transformation-based multi-task fusion method, called transformation fusion(TF), which is implemented by recursive least squares.

730, TITLE: Centrality Based Number of Cluster Estimation in Graph Clustering

<https://doi.org/10.1109/ICASSP39728.2021.9413940>

AUTHORS: M. Shamsi and S. Beheshti

HIGHLIGHT: Here, we propose a new algorithm for estimating the number of clusters in a graph using the centrality measure.

731, TITLE: Dependence-Guided Multi-View Clustering

<https://doi.org/10.1109/ICASSP39728.2021.9414971>

AUTHORS: X. Dong, D. Wu, F. Nie, R. Wang and X. Li

HIGHLIGHT: In this paper, we propose a novel approach called dependence-guided multi-view clustering (DGMC).

732, TITLE: Improved Step-Size Schedules for Noisy Gradient Methods

<https://doi.org/10.1109/ICASSP39728.2021.9414419>

AUTHORS: S. Khirirat, X. Wang, S. Magn?sson and M. Johansson

HIGHLIGHT: This paper shows that the diminishing step-size strategies can indeed be applied for a broad class of noisy gradient methods.

733, TITLE: Respipe: Resilient Model-Distributed DNN Training at Edge Networks

<https://doi.org/10.1109/ICASSP39728.2021.9413553>

AUTHORS: P. Li, E. Koyuncu and H. Seferoglu

HIGHLIGHT: In this paper, we design ResPipe, a novel resilient model-distributed DNN training mechanism against delayed/failed workers.

734, TITLE: An Optimal Stochastic Compositional Optimization Method with Applications to Meta Learning

<https://doi.org/10.1109/ICASSP39728.2021.9414369>

AUTHORS: Y. Sun, T. Chen and W. Yin

HIGHLIGHT: This paper presents a new Stochastically Corrected Stochastic Compositional gradient method (SCSC).

735, TITLE: Decentralized Optimization on Time-Varying Directed Graphs Under Communication Constraints

<https://doi.org/10.1109/ICASSP39728.2021.9415052>

AUTHORS: Y. Chen, A. Hashemi and H. Vikalo

HIGHLIGHT: We propose a communication-efficient algorithm for decentralized convex optimization that rely on sparsification of local updates exchanged between neighboring agents in the network.

736, TITLE: Decentralized Deep Learning Using Momentum-Accelerated Consensus

<https://doi.org/10.1109/ICASSP39728.2021.9414564>

AUTHORS: A. Balu, Z. Jiang, S. Y. Tan, C. Hedge, Y. M. Lee and S. Sarkar

HIGHLIGHT: In this context, we propose and analyze a novel decentralized deep learning algorithm where the agents interact over a fixed communication topology (without a central server).

737, TITLE: Demystifying Model Averaging for Communication-Efficient Federated Matrix Factorization

<https://doi.org/10.1109/ICASSP39728.2021.9413927>

AUTHORS: S. Wang, R. C. Suwandi and T. -H. Chang

HIGHLIGHT: In this paper, we investigate the federated MF problem and propose a new MA based algorithm, named FedMAvg, by judiciously combining the alternating minimization technique and MA.

738, TITLE: Sample Efficient Subspace-Based Representations for Nonlinear Meta-Learning

<https://doi.org/10.1109/ICASSP39728.2021.9414359>

AUTHORS: H. I. Gulluk, Y. Sun, S. Oymak and M. Fazel

HIGHLIGHT: This work explores a more general class of nonlinear tasks with applications ranging from binary classification, generalized linear models and neural nets.

739, TITLE: Multi-Task Learning Via Sharing Inexact Low-Rank Subspace

<https://doi.org/10.1109/ICASSP39728.2021.9414782>

AUTHORS: X. Wang and F. Nie

HIGHLIGHT: In this paper, we study the relations among multiple tasks by properly learning their shared common subspace.

740, TITLE: On The Adversarial Robustness of Principal Component Analysis

<https://doi.org/10.1109/ICASSP39728.2021.9413607>

AUTHORS: Y. Li, F. Li, L. Lai and J. Wu

HIGHLIGHT: In this paper, we investigate the adversarial robustness of principal component analysis (PCA) algorithms.

- 741, TITLE: Fast Manifold Landmarking Using Extreme Eigen-Pairs  
<https://doi.org/10.1109/ICASSP39728.2021.9414679>  
 AUTHORS: F. Wang, G. Cheung, Y. Wang and W. -T. Tan  
 HIGHLIGHT: In this paper, we select landmarks to minimize the condition number (?)
- 742, TITLE: Affine Projection Subspace Tracking  
<https://doi.org/10.1109/ICASSP39728.2021.9415032>  
 AUTHORS: M. Vil?, C. A. L?pez and J. Riba  
 HIGHLIGHT: In this paper, we consider the problem of estimating and tracking an R-dimensional subspace with relevant information embedded in an N-dimensional ambient space, given that  $N \gg R$ .
- 743, TITLE: A Hierarchical Subspace Model for Language-Attuned Acoustic Unit Discovery  
<https://doi.org/10.1109/ICASSP39728.2021.9414899>  
 AUTHORS: B. Yusuf, L. Ondel, L. Burget, J. Cernock? and M. Sara?lar  
 HIGHLIGHT: In this work, we propose a hierarchical subspace model for acoustic unit discovery.
- 744, TITLE: Independent Vector Analysis Using Semi-Parametric Density Estimation via Multivariate Entropy Maximization  
<https://doi.org/10.1109/ICASSP39728.2021.9414839>  
 AUTHORS: L. P. Damasceno, C. C. Cavalcante, T. Adali and Z. Boukouvalas  
 HIGHLIGHT: In this work, we propose a new flexible and efficient multivariate PDF estimation technique based on the maximum entropy principle and apply this technique to the development of an effective IVA algorithm that successfully matches multivariate latent sources from a wide range of distributions.
- 745, TITLE: ICA with Orthogonality Constraint: Identifiability And A New Efficient Algorithm  
<https://doi.org/10.1109/ICASSP39728.2021.9415059>  
 AUTHORS: B. Gabrielson, M. A. B. S. Akhonda, Z. Boukouvalas, S. -J. Kim and T. Adali  
 HIGHLIGHT: In this paper, we derive the identifiability conditions, starting from the orthogonal ICA maximum likelihood cost function.
- 746, TITLE: Blind Extraction of Moving Sources via Independent Component and Vector Analysis: Examples  
<https://doi.org/10.1109/ICASSP39728.2021.9413422>  
 AUTHORS: N. Amor, J. Cmejla, V. Kautsk?, Z. Koldovsk? and T. Kounovsk?  
 HIGHLIGHT: This paper is devoted to the recently proposed mixing model with constant separating vector (CSV) for Blind Source Extraction of moving sources using the FastDIVA algorithm, which is an extension of the famous FastICA and FastIVA for static mixtures.
- 747, TITLE: Single Channel Voice Separation for Unknown Number of Speakers Under Reverberant and Noisy Settings  
<https://doi.org/10.1109/ICASSP39728.2021.9413627>  
 AUTHORS: S. E. Chazan, L. Wolf, E. Nachmani and Y. Adi  
 HIGHLIGHT: We present a unified network for voice separation of an unknown number of speakers.
- 748, TITLE: Unsupervised Musical Timbre Transfer for Notification Sounds  
<https://doi.org/10.1109/ICASSP39728.2021.9414760>  
 AUTHORS: J. Yang, T. Cinquin and G. S?r?s  
 HIGHLIGHT: We present a method to transform artificial notification sounds into various musical timbres.
- 749, TITLE: Visual Privacy Protection via Mapping Distortion  
<https://doi.org/10.1109/ICASSP39728.2021.9414149>  
 AUTHORS: Y. Li, P. Liu, Y. Jiang and S. -T. Xia  
 HIGHLIGHT: In this paper, we propose the mapping distortion based protection (MDP) and its augmentation-based extension (AugMDP) to protect the data privacy by modifying the original dataset.
- 750, TITLE: L-Red: Efficient Post-Training Detection of Imperceptible Backdoor Attacks Without Access to the Training Set  
<https://doi.org/10.1109/ICASSP39728.2021.9414562>  
 AUTHORS: Z. Xiang, D. J. Miller and G. Kesidis  
 HIGHLIGHT: In this paper, we propose a Lagrangian-based RED (L-RED) that does not require knowledge of the number of source classes (or whether an attack is present).

- 751, TITLE: Multi-View Contrastive Learning for Online Knowledge Distillation  
<https://doi.org/10.1109/ICASSP39728.2021.9414664>  
AUTHORS: C. Yang, Z. An and Y. Xu  
HIGHLIGHT: We there-fore propose Multi-view Contrastive Learning (MCL) for OKD to implicitly capture correlations of feature embeddings encoded by multiple peer networks, which provide various views for understanding the input data instances.
- 752, TITLE: Dynamic Texture Recognition via Nuclear Distances on Kernelized Scattering Histogram Spaces  
<https://doi.org/10.1109/ICASSP39728.2021.9414783>  
AUTHORS: A. Sagel, J. W?rmann and H. Shen  
HIGHLIGHT: Based on the conjecture that the most distinctive characteristic of a dynamic texture is the appearance of its individual frames, this work proposes to describe dynamic textures as kernelized spaces of frame-wise feature vectors computed using the Scattering transform.
- 753, TITLE: Clustering A Collection of Networks With Mixtures of L1-Sparse Graphical Models  
<https://doi.org/10.1109/ICASSP39728.2021.9415068>  
AUTHORS: Z. Yue and V. Solo  
HIGHLIGHT: Here we consider a mixture model of sparse Gaussian graphical models and develop an exact EM algorithm that improves considerably on a previous approximation.
- 754, TITLE: Taking A Closer Look at Synthesis: Fine-Grained Attribute Analysis for Person Re-Identification  
<https://doi.org/10.1109/ICASSP39728.2021.9413757>  
AUTHORS: S. Xiang, Y. Fu, G. You and T. Liu  
HIGHLIGHT: To facilitate development in this field, we reviewed the previously developed synthetic dataset GPR and built an improved one (GPR+) with larger number of identities and distinguished attributes.
- 755, TITLE: Unified Clustering and Outlier Detection on Specialized Hardware  
<https://doi.org/10.1109/ICASSP39728.2021.9413754>  
AUTHORS: E. Cohen, H. Ushijima-Mwesigwa, A. Mandal and A. Roy  
HIGHLIGHT: In this work, we present a novel QUBO formulation of the unified clustering and outlier detection problem and use the Fujitsu Digital Annealer, a specialized CMOS hardware, to solve it.
- 756, TITLE: Class-Imbalanced Classifiers Using Ensembles of Gaussian Processes And Gaussian Process Latent Variable Models  
<https://doi.org/10.1109/ICASSP39728.2021.9414754>  
AUTHORS: L. Yang, C. Heiselman, J. G. Quirk and P. M. Djuric  
HIGHLIGHT: In this paper, binary classifiers based on Gaussian processes are chosen as bases for inferring the predictive distributions of test latent variables.
- 757, TITLE: Improving Deep Learning Sound Events Classifiers Using Gram Matrix Feature-Wise Correlations  
<https://doi.org/10.1109/ICASSP39728.2021.9414168>  
AUTHORS: A. J. Neto, A. G. C. Pacheco and D. C. Luvizon  
HIGHLIGHT: In this paper, we propose a new Sound Event Classification (SEC) method which is inspired in recent works for out-of-distribution detection.
- 758, TITLE: Adversarially Robust Classification Based on GLRT  
<https://doi.org/10.1109/ICASSP39728.2021.9413587>  
AUTHORS: B. Puranik, U. Madhow and R. Pedarsani  
HIGHLIGHT: In this paper, we explore, in the setting of classical composite hypothesis testing, a defense strategy based on the generalized likelihood ratio test (GLRT), which jointly estimates the class of interest and the adversarial perturbation.
- 759, TITLE: Cross-Corpus Speech Emotion Recognition Using Joint Distribution Adaptive Regression  
<https://doi.org/10.1109/ICASSP39728.2021.9414372>  
AUTHORS: J. Zhang, L. Jiang, Y. Zong, W. Zheng and L. Zhao  
HIGHLIGHT: In this paper, we focus on the research of cross-corpus speech emotion recognition (SER), in which the training and testing speech signals in cross-corpus SER belong to dierent speech corpus.
- 760, TITLE: Meta-Cognition-Based Simple And Effective Approach To Object Detection  
<https://doi.org/10.1109/ICASSP39728.2021.9414737>  
AUTHORS: S. P. Kumar, C. Gautam and S. Sundaram

**HIGHLIGHT:** In this paper, we explore a meta-cognitive learning strategy for object detection to improve generalization ability while at the same time maintaining detection speed.

761, **TITLE:** Graphnet: Graph Clustering with Deep Neural Networks

<https://doi.org/10.1109/ICASSP39728.2021.9413809>

**AUTHORS:** X. Zhang, J. Mu, H. Liu and X. Zhang

**HIGHLIGHT:** In this paper, we propose a novel deep graph clustering framework to tackle these two issues.

762, **TITLE:** Attention Enhanced Spatial Temporal Neural Network For HRRP Recognition

<https://doi.org/10.1109/ICASSP39728.2021.9413416>

**AUTHORS:** Y. Chu and Z. Guo

**HIGHLIGHT:** In this work, we propose a novel Attention Enhanced Convolutional Gated Recurrent Unit network (AC-GRU) for HRRP recognition which improves the representation of the spatial and temporal co-occurrence in the HRRP sequences.

763, **TITLE:** DHCN: Deep Hierarchical Context Networks For Image Annotation

<https://doi.org/10.1109/ICASSP39728.2021.9413972>

**AUTHORS:** M. Jiu and H. Sahbi

**HIGHLIGHT:** We introduce in this paper DHCN: a novel Deep Hierarchical Context Network that leverages different sources of contexts including geometric and semantic relationships.

764, **TITLE:** Online Classification of Dynamic Multilayer-Network Time Series in Riemannian Manifolds

<https://doi.org/10.1109/ICASSP39728.2021.9413560>

**AUTHORS:** C. Ye, K. Slavakis, J. Nakuci, S. F. Muldoon and J. Medaglia

**HIGHLIGHT:** This work exploits Riemannian manifolds to introduce a geometric framework for online state and community classification in dynamic multilayer networks where nodes are annotated with time series.

765, **TITLE:** On The Power of Deep But Naive Partial Label Learning

<https://doi.org/10.1109/ICASSP39728.2021.9414927>

**AUTHORS:** J. Seo and J. S. Huh

**HIGHLIGHT:** In this paper, we challenge this view by revealing the hidden power of the oldest and naivest PLL method when it is instantiated with deep neural networks.

766, **TITLE:** Advances in Morphological Neural Networks: Training, Pruning and Enforcing Shape Constraints

<https://doi.org/10.1109/ICASSP39728.2021.9415123>

**AUTHORS:** N. Dimitriadis and P. Maragos

**HIGHLIGHT:** In this paper, we study an emerging class of neural networks, the Morphological Neural networks, from some modern perspectives.

767, **TITLE:** Adversarial Learning via Probabilistic Proximity Analysis

<https://doi.org/10.1109/ICASSP39728.2021.9414096>

**AUTHORS:** J. Hollis, J. Kim and R. Raich

**HIGHLIGHT:** We propose a model-agnostic defense approach wherein the true class label of the falsified instance is inferred by analyzing its proximity to each class as measured based on class-conditional data distributions.

768, **TITLE:** Class Aware Robust Training

<https://doi.org/10.1109/ICASSP39728.2021.9414516>

**AUTHORS:** Z. Xia, B. Chen, T. Dai and S. -T. Xia

**HIGHLIGHT:** To improve robust accuracy of the important class(es), we are the first to propose a novel adversarial training method with class imbalance taken into account.

769, **TITLE:** Non-Singular Adversarial Robustness of Neural Networks

<https://doi.org/10.1109/ICASSP39728.2021.9414325>

**AUTHORS:** Y. -L. Tsai, C. -Y. Hsu, C. -M. Yu and P. -Y. Chen

**HIGHLIGHT:** In this paper, we formalize the notion of non-singular adversarial robustness for neural networks through the lens of joint perturbations to data inputs as well as model weights.

770, **TITLE:** Towards Adversarial Robustness Via Compact Feature Representations

<https://doi.org/10.1109/ICASSP39728.2021.9414696>

**AUTHORS:** M. A. Shah, R. Olivier and B. Raj



**HIGHLIGHT:** In this paper we explore hypothesis that reducing the size of the model's feature representation while maintaining its generalizability would discard spurious features while retaining perceptually relevant ones.

771, **TITLE:** Adversarial Examples Detection Beyond Image Space

<https://doi.org/10.1109/ICASSP39728.2021.9414008>

**AUTHORS:** K. Chen et al.

**HIGHLIGHT:** To detect both few-perturbation attacks and large-perturbation attacks, we propose a method beyond image space by a two-stream architecture, in which the image stream focuses on the pixel artifacts and the gradient stream copes with the confidence artifacts.

772, **TITLE:** Strong Data Augmentation Sanitizes Poisoning and Backdoor Attacks Without an Accuracy Tradeoff

<https://doi.org/10.1109/ICASSP39728.2021.9414862>

**AUTHORS:** E. Borgnia et al.

**HIGHLIGHT:** However, we find that strong data augmentations, such as mixup and CutMix, can significantly diminish the threat of poisoning and backdoor attacks without trading off performance.

773, **TITLE:** Contrastive Predictive Coding Supported Factorized Variational Autoencoder For Unsupervised Learning Of Disentangled Speech Representations

<https://doi.org/10.1109/ICASSP39728.2021.9414487>

**AUTHORS:** J. Ebberts, M. Kuhlmann, T. Cord-Landwehr and R. Haeb-Umbach

**HIGHLIGHT:** In this work we address disentanglement of style and content in speech signals.

774, **TITLE:** Contrastive Separative Coding for Self-Supervised Representation Learning

<https://doi.org/10.1109/ICASSP39728.2021.9414352>

**AUTHORS:** J. Wang, M. W. Y. Lam, D. Su and D. Yu

**HIGHLIGHT:** To extract robust deep representations from long sequential modeling of speech data, we propose a self-supervised learning approach, namely Contrastive Separative Coding (CSC).

775, **TITLE:** Contrastive Semi-Supervised Learning for ASR

<https://doi.org/10.1109/ICASSP39728.2021.9414079>

**AUTHORS:** A. Xiao, C. Fuegen and A. Mohamed

**HIGHLIGHT:** Inspired by the successes of contrastive representation learning for both computer vision and speech applications, and more recently for supervised learning of visual objects [1], we propose Contrastive Semi-supervised Learning (CSL).

776, **TITLE:** Contrastive Learning of General-Purpose Audio Representations

<https://doi.org/10.1109/ICASSP39728.2021.9413528>

**AUTHORS:** A. Saeed, D. Grangier and N. Zeghidour

**HIGHLIGHT:** We introduce COLA, a self-supervised pre-training approach for learning a general-purpose representation of audio.

777, **TITLE:** SEQ-CPC : Sequential Contrastive Predictive Coding for Automatic Speech Recognition

<https://doi.org/10.1109/ICASSP39728.2021.9413645>

**AUTHORS:** Y. Chen et al.

**HIGHLIGHT:** Inspired by the contrastive predictive coding (CPC), we propose a feature representation scheme for automatic speech recognition (ASR), which encodes sequential dependency information from raw audio signals.

778, **TITLE:** On Scaling Contrastive Representations for Low-Resource Speech Recognition

<https://doi.org/10.1109/ICASSP39728.2021.9414310>

**AUTHORS:** L. Borgholt, T. M. S. Tax, J. D. Havtorn, L. Maal'e and C. Igel

**HIGHLIGHT:** We explore the performance of such systems without fine-tuning by training a state-of-the-art speech recognizer on the fixed representations from the computationally demanding wav2vec 2.0 framework.

779, **TITLE:** Convex Neural Autoregressive Models: Towards Tractable, Expressive, and Theoretically-Backed Models for Sequential Forecasting and Generation

<https://doi.org/10.1109/ICASSP39728.2021.9413662>

**AUTHORS:** V. Gupta, B. Bartan, T. Ergen and M. Pilanci

**HIGHLIGHT:** Consequently, we introduce techniques to derive tractable, expressive, and theoretically-interpretable models that are nearly equivalent to neural autoregressive models.

780, **TITLE:** Inertial Proximal Deep Learning Alternating Minimization for Efficient Neural Network Training

<https://doi.org/10.1109/ICASSP39728.2021.9413500>

AUTHORS: L. Qiao, T. Sun, H. Pan and D. Li

HIGHLIGHT: This work develops an improved DLAM by the well-known inertial technique, namely iPDLAM, which predicts a point by linearization of current and last iterates.

781, TITLE: Kalman Optimizer for Consistent Gradient Descent

<https://doi.org/10.1109/ICASSP39728.2021.9414588>

AUTHORS: X. Yang

HIGHLIGHT: In this paper, we propose Kalman Optimizer (KO), an efficient stochastic optimization algorithm that adopts Kalman filter to make consistent estimation of the local gradient by solving an adaptive filtering problem.

782, TITLE: Kalmannet: Data-Driven Kalman Filtering

<https://doi.org/10.1109/ICASSP39728.2021.9413750>

AUTHORS: G. Revach, N. Shlezinger, R. J. G. van Sloun and Y. C. Eldar

HIGHLIGHT: We present an offline training method, and numerically illustrate that KalmanNet can achieve optimal performance without full knowledge of the model parameters.

783, TITLE: HCGM-Net: A Deep Unfolding Network for Financial Index Tracking

<https://doi.org/10.1109/ICASSP39728.2021.9414430>

AUTHORS: R. Pauwels, E. Tsiliogianni and N. Deligiannis

HIGHLIGHT: In this paper, we focus on sparse index tracking and employ a Frank-Wolfe-based algorithm which we translate into a deep neural network, a strategy known as deep unfolding.

784, TITLE: Augmenting Transferred Representations for Stock Classification

<https://doi.org/10.1109/ICASSP39728.2021.9413530>

AUTHORS: E. Fons, P. Dawson, X. -j. Zeng, J. Keane and A. Iosifidis

HIGHLIGHT: In this paper we show that using transfer learning can help with this task, by pre-training a model to extract universal features on the full universe of stocks of the S&P500 index and then transferring it to another model to directly learn a trading rule.

785, TITLE: A Framework for Pruning Deep Neural Networks Using Energy-Based Models

<https://doi.org/10.1109/ICASSP39728.2021.9414434>

AUTHORS: H. Salehinejad and S. Valaee

HIGHLIGHT: In this paper, we propose a framework for pruning DNNs based on a population-based global optimization method.

786, TITLE: Prototype-Based Personalized Pruning

<https://doi.org/10.1109/ICASSP39728.2021.9414526>

AUTHORS: J. Kim, S. Chang, S. Yun and N. Kwak

HIGHLIGHT: In this work, we propose a dynamic personalization method called prototype-based personalized pruning (PPP).

787, TITLE: Tensor Reordering for CNN Compression

<https://doi.org/10.1109/ICASSP39728.2021.9413944>

AUTHORS: M. Ulicny, V. A. Krylov and R. Dahyot

HIGHLIGHT: We show how parameter redundancy in Convolutional Neural Network (CNN) filters can be effectively reduced by pruning in spectral domain.

788, TITLE: Pruning of Convolutional Neural Networks using Ising Energy Model

<https://doi.org/10.1109/ICASSP39728.2021.9414645>

AUTHORS: H. Salehinejad and S. Valaee

HIGHLIGHT: In this paper, we propose an Ising energy model within an optimization framework for pruning convolutional kernels and hidden units.

789, TITLE: Reweighted Dynamic Group Convolution

<https://doi.org/10.1109/ICASSP39728.2021.9414570>

AUTHORS: W. Chen, C. Wang, Z. Zhang, Z. Huo and L. Gao

HIGHLIGHT: Inspired by the previous work, a new reweighted dynamic group convolution (RDGC) structure, including a reweighted pruning module and a survival loss, is proposed in this work for more precise channel pruning.

790, TITLE: Layer-Wise Interpretation of Deep Neural Networks using Identity Initialization

<https://doi.org/10.1109/ICASSP39728.2021.9414873>

AUTHORS: S. Kubota, H. Hayashi, T. Hayase and S. Uchida

HIGHLIGHT: In this paper, we propose an interpretation method for a deep multilayer perceptron, which is the most general architecture of NNs, based on identity initialization (namely, initialization using identity matrices).

791, TITLE: Detection of Post-Traumatic Stress Disorder Using Learned Time-Frequency Representations from Pupillometry

<https://doi.org/10.1109/ICASSP39728.2021.9414988>

AUTHORS: B. Taha, M. Kirk, P. Ritvo and D. Hatzinakos

HIGHLIGHT: This study investigates the utility of pupillometry as a biomarker to detect PTSD in a sample of 39 adults with (n = 22) and without (n = 17) PTSD.

792, TITLE: Subject-Invariant Eeg Representation Learning For Emotion Recognition

<https://doi.org/10.1109/ICASSP39728.2021.9414496>

AUTHORS: S. Rayatdoost, Y. Yin, D. Rudrauf and M. Soleymani

HIGHLIGHT: In this paper, we propose an adversarial deep domain adaptation approach for emotion recognition from electroencephalogram (EEG) signals.

793, TITLE: Towards Parkinson's Disease Prognosis Using Self-Supervised Learning and Anomaly Detection

<https://doi.org/10.1109/ICASSP39728.2021.9414840>

AUTHORS: H. Jiang, W. Y. Bryan Lim, J. Shyuan Ng, Y. Wang, Y. Chi and C. Miao

HIGHLIGHT: We propose to tackle the problem as a semi-supervised anomaly detection task, where we model the physiological patterns of healthy subjects instead.

794, TITLE: In-Bed Pressure-Based Pose Estimation Using Image Space Representation Learning

<https://doi.org/10.1109/ICASSP39728.2021.9413516>

AUTHORS: V. Davoodnia, S. Ghorbani and A. Etemad

HIGHLIGHT: In this paper, we address this challenge by presenting a novel end-to-end framework capable of accurately locating body parts from vague pressure data.

795, TITLE: Towards The Development of Subject-Independent Inverse Metabolic Models

<https://doi.org/10.1109/ICASSP39728.2021.9413829>

AUTHORS: S. Sajjadi et al.

HIGHLIGHT: To address this issue, we describe an approach to monitor diet automatically, by analyzing fluctuations in glucose after a meal is consumed.

796, TITLE: Human-Expert-Level Brain Tumor Detection Using Deep Learning with Data Distillation And Augmentation

<https://doi.org/10.1109/ICASSP39728.2021.9415067>

AUTHORS: D. Lu, N. Polomac, I. Gacheva, E. Hattingen and J. Triesch

HIGHLIGHT: To overcome these challenges, we propose a new method for training a deep neural network that distills particularly representative training examples and augments the training data by mixing these samples from one class with those from the same and other classes to create additional training samples.

797, TITLE: Multimodal Punctuation Prediction with Contextual Dropout

<https://doi.org/10.1109/ICASSP39728.2021.9414979>

AUTHORS: A. Silva, B. -J. Theobald and N. Apostoloff

HIGHLIGHT: We first present a transformer-based approach for punctuation prediction that achieves 8% improvement on the IWSLT 2012 TED Task, beating the previous state of the art [1]. We next describe our multimodal model that learns from both text and audio, which achieves 8% improvement over the text-only algorithm on an internal dataset for which we have both the audio and transcriptions.

798, TITLE: Multi-Modal Label Dequantized Gaussian Process Latent Variable Model for Ordinal Label Estimation

<https://doi.org/10.1109/ICASSP39728.2021.9415090>

AUTHORS: M. Matsumoto, K. Maeda, N. Saito, T. Ogawa and M. Haseyama

HIGHLIGHT: This paper presents multi-modal label dequantized Gaussian process latent variable model (mLDGP) for ordinal label estimation.

799, TITLE: Generative Information Fusion

<https://doi.org/10.1109/ICASSP39728.2021.9414284>

AUTHORS: K. Tran, W. Sakla and H. Krim

**HIGHLIGHT:** In this work, we demonstrate the ability to exploit sensing modalities for mitigating an unrepresented modality or for potentially re-targeting resources.

800, **TITLE:** Self-Augmented Multi-Modal Feature Embedding

<https://doi.org/10.1109/ICASSP39728.2021.9413974>

**AUTHORS:** S. Matsuo, S. Uchida and B. K. Iwana

**HIGHLIGHT:** To exploit this fact, we propose the use of self-augmentation and combine it with multi-modal feature embedding.

801, **TITLE:** Optimize What Matters: Training DNN-Hmm Keyword Spotting Model Using End Metric

<https://doi.org/10.1109/ICASSP39728.2021.9414797>

**AUTHORS:** A. Shrivastava, A. Kundu, C. Dhir, D. Naik and O. Tuzel

**HIGHLIGHT:** We address this loss-metric mismatch with a novel end-to-end training strategy that learns the DNN parameters by optimizing for the detection score.

802, **TITLE:** Co-Attentional Transformers for Story-Based Video Understanding

<https://doi.org/10.1109/ICASSP39728.2021.9413868>

**AUTHORS:** B. Bebensee and B. -T. Zhang

**HIGHLIGHT:** We propose a novel co-attentional transformer model to better capture long-term dependencies seen in visual stories such as dramas and measure its performance on the video question answering task.

803, **TITLE:** Deep Generative Demixing: Error Bounds for Demixing Subgaussian Mixtures of Lipschitz Signals

<https://doi.org/10.1109/ICASSP39728.2021.9413573>

**AUTHORS:** A. Berk

**HIGHLIGHT:** Here, we investigate the subgaussian demixing problem for two Lipschitz signals, with GNN demixing as a special case.

804, **TITLE:** Towards an Intrinsic Definition of Robustness for a Classifier

<https://doi.org/10.1109/ICASSP39728.2021.9414573>

**AUTHORS:** T. Giraudon, V. Gripon, M. L?we and F. Vermet

**HIGHLIGHT:** In this paper, we point out that averaging the radius of robustness of samples in a validation set is a statistically weak measure.

805, **TITLE:** Phase Transitions for One-Vs-One and One-Vs-All Linear Separability in Multiclass Gaussian Mixtures

<https://doi.org/10.1109/ICASSP39728.2021.9414099>

**AUTHORS:** G. R. Kini and C. Thrampoulidis

**HIGHLIGHT:** We present precise formulae characterizing the phase transitions as a function of the data geometry and the number of classes.

806, **TITLE:** Leaky Integrator Dynamical Systems and Reachable Sets

<https://doi.org/10.1109/ICASSP39728.2021.9413667>

**AUTHORS:** B. Whiteaker and P. Gerstoft

**HIGHLIGHT:** This work applies the controllability matrix of control theory to quickly identify a reduced size replacement reservoir.

807, **TITLE:** Benign Overfitting in Binary Classification of Gaussian Mixtures

<https://doi.org/10.1109/ICASSP39728.2021.9413946>

**AUTHORS:** K. Wang and C. Thrampoulidis

**HIGHLIGHT:** This paper studies benign-overfitting for data generated from a popular binary Gaussian mixtures model (GMM) and classifiers trained by support-vector machines (SVM).

808, **TITLE:** An Order-Optimal Adaptive Test Plan for Noisy Group Testing Under Unknown Noise Models

<https://doi.org/10.1109/ICASSP39728.2021.9414111>

**AUTHORS:** S. Salgia and Q. Zhao

**HIGHLIGHT:** We propose an adaptive test plan consisting of a hierarchy of biased random walks guided by a local sequential test which together lend adaptivity and agnosticism to the unknown noise model.

809, **TITLE:** SapAugment: Learning A Sample Adaptive Policy for Data Augmentation

<https://doi.org/10.1109/ICASSP39728.2021.9413928>

**AUTHORS:** T. -Y. Hu et al.

**HIGHLIGHT:** To formalize these intuitions, we propose a novel method to learn a Sample-Adaptive Policy for Augmentation ? SapAugment.

810, **TITLE:** Hierarchical Coded Elastic Computing

<https://doi.org/10.1109/ICASSP39728.2021.9414863>

**AUTHORS:** S. Kiani, T. Adikari and S. C. Draper

**HIGHLIGHT:** In this paper, we propose two hierarchical coded elastic computing schemes that can further speed up the system by exploiting stragglers and effectively allocating tasks among available nodes.

811, **TITLE:** Synthesize & Learn: Jointly Optimizing Generative and Classifier Networks for Improved Drowsiness Detection

<https://doi.org/10.1109/ICASSP39728.2021.9413822>

**AUTHORS:** S. Banerjee, A. Joshi, A. Ghoneim, S. Kyal and T. Mishra

**HIGHLIGHT:** We focus on the problem of alleviating the class imbalance problem by using generative adversarial networks (GAN) to synthesize examples of sparse classes directly in the feature-space.

812, **TITLE:** A Joint Convolutional and Spatial Quad-Directional LSTM Network for Phase Unwrapping

<https://doi.org/10.1109/ICASSP39728.2021.9414748>

**AUTHORS:** M. V. Perera and A. De Silva

**HIGHLIGHT:** In this paper, we introduce a novel Convolutional Neural Network (CNN) that incorporates a Spatial Quad-Directional Long Short Term Memory (SQD-LSTM) for phase unwrapping, by formulating it as a regression problem.

813, **TITLE:** Integrated Classification and Localization of Targets Using Bayesian Framework In Automotive Radars

<https://doi.org/10.1109/ICASSP39728.2021.9414131>

**AUTHORS:** A. Dubey, A. Santra, J. Fuchs, M. L?bke, R. Weigel and F. Lurz

**HIGHLIGHT:** Compared to state-of-the-art using independent classification and tracking, in this paper, we propose an integrated tracker and classifier leading to a novel Bayesian framework.

814, **TITLE:** A DNN Autoencoder for Automotive Radar Interference Mitigation

<https://doi.org/10.1109/ICASSP39728.2021.9413619>

**AUTHORS:** S. Chen, J. Taghia, T. Fei, U. K?hnau, N. Pohl and R. Martin

**HIGHLIGHT:** In this paper, a novel interference mitigation approach using an autoencoder in combination with a traditional interference detection filter is introduced.

815, **TITLE:** DURAS: Deep Unfolded Radar Sensing Using Doppler Focusing

<https://doi.org/10.1109/ICASSP39728.2021.9414967>

**AUTHORS:** P. Goyal, S. Mulleti, A. Gupta and Y. C. Eldar

**HIGHLIGHT:** In this paper, we propose Deep Unfolded Radar Sensing (DURAS), a model-based deep learning architecture to address this problem.

816, **TITLE:** NNAKF: A Neural Network Adapted Kalman Filter for Target Tracking

<https://doi.org/10.1109/ICASSP39728.2021.9414681>

**AUTHORS:** S. Jouaber, S. Bonnabel, S. Velasco-Forero and M. Pilt?

**HIGHLIGHT:** In the present paper we advocate a novel method to increase Q during maneuvers (and hence the Kalman gains) based on a recurrent neural network (RNN).

817, **TITLE:** Multi-Sample Online Learning for Spiking Neural Networks Based on Generalized Expectation Maximization

<https://doi.org/10.1109/ICASSP39728.2021.9414804>

**AUTHORS:** H. Jang and O. Simeone

**HIGHLIGHT:** While prior work used single-sample estimators obtained from a single run of the network, this paper proposes to leverage multiple compartments that sample independent spiking signals while sharing synaptic weights.

818, **TITLE:** Probabilistic Graph Neural Networks for Traffic Signal Control

<https://doi.org/10.1109/ICASSP39728.2021.9414829>

**AUTHORS:** T. Zhong, Z. Xu and F. Zhou

**HIGHLIGHT:** This work presents a variational graph learning model TSC-GNN (Traffic Signal Control via probabilistic Graph Neural Networks) to learn the latent representations of agents and generate Q-value while taking traffic uncertainty conditions into account.

819, **TITLE:** Task-Aware Neural Architecture Search

<https://doi.org/10.1109/ICASSP39728.2021.9414412>

AUTHORS: C. P. Le, M. Soltani, R. Ravier and V. Tarokh

HIGHLIGHT: In this paper, we propose a novel framework for neural architecture search, utilizing a dictionary of models of base tasks and the similarity between the target task and the atoms of the dictionary; hence, generating an adaptive search space based on the base models of the dictionary.

820, TITLE: F-Net: Fusion Neural Network for Vehicle Trajectory Prediction in Autonomous Driving

<https://doi.org/10.1109/ICASSP39728.2021.9413881>

AUTHORS: J. Wang, P. Wang, C. Zhang, K. Su and J. Li

HIGHLIGHT: In this paper, based on recurrent neural networks and convolutional neural networks, we propose a fusion neural network architecture named F-Net to deal with vehicle trajectory prediction on highway and urban scenarios in autonomous driving applications.

821, TITLE: Unsupervised Reconstruction of Sea Surface Currents from AIS Maritime Traffic Data Using Learnable Variational Models

<https://doi.org/10.1109/ICASSP39728.2021.9415038>

AUTHORS: S. Bena?chouche, C. Le Goff, Y. Guichoux, F. Rousseau and R. Fablet

HIGHLIGHT: Here, we show that an unsupervised variational learning scheme provides new means to elucidate how AIS data streams can be converted into sea surface currents.

822, TITLE: A Compact Joint Distillation Network for Visual Food Recognition

<https://doi.org/10.1109/ICASSP39728.2021.9415019>

AUTHORS: Z. Heng, K. -H. Yap and A. C. Kot

HIGHLIGHT: In view of this, this paper proposes a joint distillation framework that targets to achieve a high visual food recognition accuracy using a compact network.

823, TITLE: Pipeline Safety Early Warning Method for Distributed Signal using Bilinear CNN and LightGBM

<https://doi.org/10.1109/ICASSP39728.2021.9414544>

AUTHORS: Y. Yang, Y. Li and H. Zhang

HIGHLIGHT: In this study, we utilized a novel machine learning method based on the spatiotemporal features of distributed optical fiber sensor signals to monitor the safety of oil and gas pipelines in real time.

824, TITLE: Deep Learning Based Hybrid Precoding in Dual-Band Communication Systems

<https://doi.org/10.1109/ICASSP39728.2021.9414488>

AUTHORS: R. Ismayilov, R. L. G. Cavalcante and S. Stanczak

HIGHLIGHT: We propose a deep learning-based method that uses spatial and temporal information extracted from the sub-6GHz band to predict/track beams in the millimeter-wave (mmWave) band.

825, TITLE: Deep Learning-Based Cross-Layer Resource Allocation for Wired Communication Systems

<https://doi.org/10.1109/ICASSP39728.2021.9413777>

AUTHORS: P. Behmandpoor, J. Verdyck and M. Moonen

HIGHLIGHT: In this paper, a cross-layer resource allocation (RA) scheme based on deep learning is introduced for multi-tone multi-user wired communication systems such as, for instance, digital subscriber line (DSL) systems under the current G.fast standard.

826, TITLE: ATVIO: Attention Guided Visual-Inertial Odometry

<https://doi.org/10.1109/ICASSP39728.2021.9413912>

AUTHORS: L. Liu, G. Li and T. H. Li

HIGHLIGHT: In the paper, we propose a novel attention guided deep framework for visual-inertial odometry (ATVIO) to improve the performance of VIO.

827, TITLE: Feature Integration via Semi-Supervised Ordinally Multi-Modal Gaussian Process Latent Variable Model

<https://doi.org/10.1109/ICASSP39728.2021.9414109>

AUTHORS: K. Kamikawa, K. Maeda, T. Ogawa and M. Haseyama

HIGHLIGHT: This paper presents a method of feature integration via semi-supervised ordinally multi-modal Gaussian process latent variable model (Semi-OMGP).

828, TITLE: A Multi-Layer Multi-Channel Attentive Network for Gender and Age Recognition

<https://doi.org/10.1109/ICASSP39728.2021.9413913>

AUTHORS: J. Chen, H. Yu and Y. Kang

HIGHLIGHT: In this paper a multi-layer multi-channel attentive network based on the idea of divide-and-conquer is proposed.

- 829, TITLE: Effect of Language Proficiency on Subjective Evaluation of Noise Suppression Algorithms  
<https://doi.org/10.1109/ICASSP39728.2021.9413834>  
AUTHORS: B. Naderi, G. Mittag, R. Z. Jimenez and S. Moller  
HIGHLIGHT: Whereas standard tests for assessing perceived quality make use of native listeners, we assume that noise-reduced speech and residual noise may affect native and non-native listeners of a target language in different ways.
- 830, TITLE: Melody Harmonization Using Orderless Nade, Chord Balancing, and Blocked Gibbs Sampling  
<https://doi.org/10.1109/ICASSP39728.2021.9414281>  
AUTHORS: C. -E. Sun, Y. -W. Chen, H. -S. Lee, Y. -H. Chen and H. -M. Wang  
HIGHLIGHT: In this study, we apply the concept of orderless NADE, which takes the melody and its partially masked chord sequence as the input of the BiLSTM-based networks to learn the masked ground truth, to the training process.
- 831, TITLE: Cross-Domain Semi-Supervised Deep Metric Learning for Image Sentiment Analysis  
<https://doi.org/10.1109/ICASSP39728.2021.9414150>  
AUTHORS: Y. Liang, K. Maeda, T. Ogawa and M. Haseyama  
HIGHLIGHT: This paper presents a novel method on image sentiment analysis called cross-domain semi-supervised deep metric learning (CDSS-DML).
- 832, TITLE: Audiovisual Highlight Detection in Videos  
<https://doi.org/10.1109/ICASSP39728.2021.9413394>  
AUTHORS: K. Mundnich, A. Fenster, A. Khare and S. Sundaram  
HIGHLIGHT: In this paper, we test the hypothesis that interesting events in unstructured videos are inherently audiovisual.
- 833, TITLE: Teacher-Assisted Mini-Batch Sampling for Blind Distillation Using Metric Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9414294>  
AUTHORS: N. Inoue  
HIGHLIGHT: The proposed framework introduces metric learning to blind distillation.
- 834, TITLE: Rule-Embedded Network for Audio-Visual Voice Activity Detection in Live Musical Video Streams  
<https://doi.org/10.1109/ICASSP39728.2021.9413418>  
AUTHORS: Y. Hou, Y. Deng, B. Zhu, Z. Ma and D. Botteldooren  
HIGHLIGHT: This paper proposes a rule-embedded network to fuse the audio-visual (A-V) inputs for better detection of the target voice.
- 835, TITLE: Reinforcement Stacked Learning with Semantic-Associated Attention for Visual Question Answering  
<https://doi.org/10.1109/ICASSP39728.2021.9414636>  
AUTHORS: X. Xiao, C. Zhang, S. Xiang and C. Pan  
HIGHLIGHT: In this paper, depending on these problems, a semantic-associated attention method and a reinforcement stacked learning mechanism are proposed.
- 836, TITLE: Hierarchical Refined Attention for Scene Text Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9413534>  
AUTHORS: M. Zhang, M. Ma and P. Wang  
HIGHLIGHT: This paper proposes a novel framework named hierarchical refined attention network (HRAN) for STR.
- 837, TITLE: Collaborative Learning to Generate Audio-Video Jointly  
<https://doi.org/10.1109/ICASSP39728.2021.9413802>  
AUTHORS: V. K. Kurmi, V. Bajaj, B. N. Patro, K. S. Venkatesh, V. P. Namboodiri and P. Jyothi  
HIGHLIGHT: Towards this, we propose a method that demonstrates that we are able to generate naturalistic samples of video and audio data by the joint correlated generation of audio and video modalities.
- 838, TITLE: An Attention-Seq2Seq Model Based on CRNN Encoding for Automatic Labanotation Generation from Motion Capture Data  
<https://doi.org/10.1109/ICASSP39728.2021.9414976>  
AUTHORS: M. Li, Z. Miao, X. -P. Zhang and W. Xu  
HIGHLIGHT: In this paper, we propose an attention-seq2seq model based on Convolutional Recurrent Neural Network (CRNN).
- 839, TITLE: Show and Speak: Directly Synthesize Spoken Description of Images



<https://doi.org/10.1109/ICASSP39728.2021.9414021>

AUTHORS: X. Wang, S. Feng, J. Zhu, M. Hasegawa-Johnson and O. Scharenborg

HIGHLIGHT: This paper proposes a new model, referred to as the show and speak (SAS) model that, for the first time, is able to directly synthesize spoken descriptions of images, bypassing the need for any text or phonemes.

840, TITLE: Drawgan: Text to Image Synthesis with Drawing Generative Adversarial Networks

<https://doi.org/10.1109/ICASSP39728.2021.9414166>

AUTHORS: Z. Zhang, J. Zhou, W. Yu and N. Jiang

HIGHLIGHT: In this paper, we propose a novel drawing generative adversarial networks (DrawGAN) for text-to-image synthesis.

841, TITLE: Disentangling Subject-Dependent/-Independent Representations for 2D Motion Retargeting

<https://doi.org/10.1109/ICASSP39728.2021.9413715>

AUTHORS: F. Xie, G. Irie and T. Matsubayashi

HIGHLIGHT: In this work, we propose a novel network to separate subject-dependent and -independent motion features and to reconstruct a new skeleton with or without subject-dependent motion features.

842, TITLE: Network and Content-Dependent Bitrate Ladder Estimation for Adaptive Bitrate Video Streaming

<https://doi.org/10.1109/ICASSP39728.2021.9413558>

AUTHORS: P. Lebreton and K. Yamagishi

HIGHLIGHT: In this paper, a method is presented to estimate bitrate ladders on the basis of both content complexity and network traces.

843, TITLE: Ultra-Low Bitrate Video Conferencing Using Deep Image Animation

<https://doi.org/10.1109/ICASSP39728.2021.9414731>

AUTHORS: G. Konuko, G. Valenzise and S. Lathuilière

HIGHLIGHT: In this work we propose a novel deep learning approach for ultra-low bitrate video compression for video conferencing applications.

844, TITLE: Hierarchical Bit-Wise Differential Coding (HBDC) of Point Cloud Attributes

<https://doi.org/10.1109/ICASSP39728.2021.9413427>

AUTHORS: Y. Huang, B. Wang, C. -C. J. Kuo, H. Yuan and J. Peng

HIGHLIGHT: Targeting both computing and coding efficiencies, we propose in this work a novel hierarchical bit-wise differential coding scheme to compress point cloud attributes.

845, TITLE: Learning-Based Lossless Compression of 3D Point Cloud Geometry

<https://doi.org/10.1109/ICASSP39728.2021.9414763>

AUTHORS: D. T. Nguyen, M. Quach, G. Valenzise and P. Duhamel

HIGHLIGHT: This paper presents a learning-based, lossless compression method for static point cloud geometry, based on context-adaptive arithmetic coding.

846, TITLE: Image Coding with Neural Network-Based Colorization

<https://doi.org/10.1109/ICASSP39728.2021.9413816>

AUTHORS: D. Lopes, J. Ascenso, C. Brites and F. Pereira

HIGHLIGHT: Motivated by the excellent results obtained with deep learning-based solutions in the area of automatic colorization, this paper proposes an image coding solution integrating a deep learning-based colorization process to estimate the chrominance components based on the decoded luminance which is regularly encoded with a conventional image coding standard.

847, TITLE: Joint Reinforcement Learning and Game Theory Bitrate Control Method for 360-Degree Dynamic Adaptive Streaming

<https://doi.org/10.1109/ICASSP39728.2021.9414370>

AUTHORS: X. Wei, M. Zhou, S. Kwong, H. Yuan and T. Xiang

HIGHLIGHT: A joint reinforcement learning (RL) and game theory method is presented for segment-level continuous bitrate selection and tile-level bitrate allocation in tile-based 360-degree streaming to increase users' quality of experience (QoE).

848, TITLE: HCAG: A Hierarchical Context-Aware Graph Attention Model for Depression Detection

<https://doi.org/10.1109/ICASSP39728.2021.9413486>

AUTHORS: M. Niu, K. Chen, Q. Chen and L. Yang

HIGHLIGHT: In this paper, we propose HCAG, a novel Hierarchical Context-Aware Graph attention model for ADD.

- 849, TITLE: When Face Recognition Meets Occlusion: A New Benchmark  
<https://doi.org/10.1109/ICASSP39728.2021.9413893>  
AUTHORS: B. Huang et al.  
HIGHLIGHT: To this end, we pioneer a simulated occlusion face recognition dataset.
- 850, TITLE: A Triplet Appearance Parsing Network for Person Re-Identification  
<https://doi.org/10.1109/ICASSP39728.2021.9413666>  
AUTHORS: M. Xiong et al.  
HIGHLIGHT: This study develops a Triplet person Appearances Parsing Framework (TAPF) which eliminates the surrounding interference factors of bounding boxes for person re-identification.
- 851, TITLE: Part-Aligned Network with Background for Misaligned Person Search  
<https://doi.org/10.1109/ICASSP39728.2021.9415107>  
AUTHORS: X. Zhong, Y. Liu, W. Huang, X. Wang, B. Ma and J. Yuan  
HIGHLIGHT: In this paper, we propose a part-aligned network with background (PANB) to address this misalignment issue.
- 852, TITLE: Learning Pose-Adaptive Lip Sync with Cascaded Temporal Convolutional Network  
<https://doi.org/10.1109/ICASSP39728.2021.9413472>  
AUTHORS: R. Zheng, B. Song and C. Ji  
HIGHLIGHT: In this paper, we design a cascaded temporal convolutional network to successively generate mouth shapes and corresponding jawlines based on audio signals and template headposes.
- 853, TITLE: Assessment of Bipolar Disorder Using Heterogeneous Data of Smartphone-Based Digital Phenotyping  
<https://doi.org/10.1109/ICASSP39728.2021.9415008>  
AUTHORS: H. -Y. Su, C. -H. Wu, C. -R. Liou, E. C. -L. Lin and P. See Chen  
HIGHLIGHT: In order to reduce the cost of social and medical resources, this study collects the user's data by the App on smartphones, consisting of location data (GPS), self-report scales, daily mood, sleeping time and records of multi-media (text, speech, video) which are heterogeneous digital phenotyping data, to build a database.
- 854, TITLE: Multi-Granularity Feature Interaction and Relation Reasoning for 3D Dense Alignment and Face Reconstruction  
<https://doi.org/10.1109/ICASSP39728.2021.9413649>  
AUTHORS: L. Li, X. Li, K. Wu, K. Lin and S. Wu  
HIGHLIGHT: In this paper, we propose a multi-granularity feature interaction and relation reasoning network (MFIRRN) which can recover a detail-rich 3D face and perform more accurate dense alignment in an unconstrained environment.
- 855, TITLE: Independent Sign Language Recognition with 3d Body, Hands, and Face Reconstruction  
<https://doi.org/10.1109/ICASSP39728.2021.9414278>  
AUTHORS: A. Kratimenos, G. Pavlakos and P. Maragos  
HIGHLIGHT: In this work, we employ SMPL-X, a contemporary parametric model that enables joint extraction of 3D body shape, face and hands information from a single image.
- 856, TITLE: Multimodal Cross- and Self-Attention Network for Speech Emotion Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9414654>  
AUTHORS: L. Sun, B. Liu, J. Tao and Z. Lian  
HIGHLIGHT: In this paper, we propose a novel Multimodal Cross- and Self-Attention Network (MCSAN) to tackle this problem.
- 857, TITLE: Multi-Target DoA Estimation with an Audio-Visual Fusion Mechanism  
<https://doi.org/10.1109/ICASSP39728.2021.9413776>  
AUTHORS: X. Qian, M. Madhavi, Z. Pan, J. Wang and H. Li  
HIGHLIGHT: With this motivation, we propose to use neural networks with audio and visual signals for multi-speaker localization.
- 858, TITLE: Improving Multimodal Speech Enhancement by Incorporating Self-Supervised and Curriculum Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9413431>  
AUTHORS: Y. Cheng, M. He, J. Yu and R. Feng  
HIGHLIGHT: In this paper, we present a co-attention based framework that incorporates self-supervised and curriculum learning to derive the target speech in noisy environments.

- 859, TITLE: Autoencoder for Vibrotactile Signal Compression  
<https://doi.org/10.1109/ICASSP39728.2021.9413370>  
AUTHORS: Z. Li, R. Hassen and Z. Wang  
HIGHLIGHT: Inspired by the recent success of deep neural network (DNN) based autoencoder, we make the first attempt to apply autoencoder for lossy compression of haptic vibrotactile signals, where a convolutional neural network (CNN) and a rate-distortion (RD) function are used as the transform and cost functions, respectively.
- 860, TITLE: Cross-Modal Knowledge Distillation For Fine-Grained One-Shot Classification  
<https://doi.org/10.1109/ICASSP39728.2021.9414480>  
AUTHORS: J. Zhao, X. Lin, Y. Yang, J. Yang and L. He  
HIGHLIGHT: In this paper, we adopt domain-specific knowledge to fill the gap of insufficient annotated data.
- 861, TITLE: Learning Audio-Visual Correlations From Variational Cross-Modal Generation  
<https://doi.org/10.1109/ICASSP39728.2021.9414296>  
AUTHORS: Y. Zhu, Y. Wu, H. Latapie, Y. Yang and Y. Yan  
HIGHLIGHT: We introduce a novel Variational AutoEncoder (VAE) framework that consists of Multiple encoders and a Shared decoder (MS-VAE) with an additional Wasserstein distance constraint to tackle the problem.
- 862, TITLE: ECCL: Explicit Correlation-Based Convolution Boundary Locator for Moment Localization  
<https://doi.org/10.1109/ICASSP39728.2021.9414047>  
AUTHORS: X. Liu, X. Nie, J. Teng, F. Hao and Y. Yin  
HIGHLIGHT: In this paper, we present a new boundary-determining strategy called explicit correlation-based convolution boundary locator (ECCL), which can handle any lengths of videos and moments while leveraging fine-grained matching relationships.
- 863, TITLE: COOPNet: Multi-Modal Cooperative Gender Prediction in Social Media User Profiling  
<https://doi.org/10.1109/ICASSP39728.2021.9414808>  
AUTHORS: L. Li, K. Hu, Y. Zheng, J. Liu and K. A. Lee  
HIGHLIGHT: In this paper, we propose a novel text-image cooperation framework (COOPNet), a bridge connection network architecture that exchanges information between texts and images.
- 864, TITLE: Robust Latent Representations Via Cross-Modal Translation and Alignment  
<https://doi.org/10.1109/ICASSP39728.2021.9413456>  
AUTHORS: V. Rajan, A. Brutti and A. Cavallaro  
HIGHLIGHT: To address this limitation, we aim to improve the testing performance of uni-modal systems using multiple modalities during training only.
- 865, TITLE: Semi-Supervised Multimodal Image Translation for Missing Modality Imputation  
<https://doi.org/10.1109/ICASSP39728.2021.9413461>  
AUTHORS: W. Sun, F. Ma, Y. Li, S. -L. Huang, S. Ni and L. Zhang  
HIGHLIGHT: We propose a semi-supervised algorithm for multimodal learning with missing data, namely Cyclic Autoencoder (CycAE).
- 866, TITLE: Deep Adversarial Quantization Network for Cross-Modal Retrieval  
<https://doi.org/10.1109/ICASSP39728.2021.9414247>  
AUTHORS: Y. Zhou, Y. Feng, M. Zhou, B. Qiang, L. Hou U and J. Zhu  
HIGHLIGHT: In this paper, we propose a seamless multimodal binary learning method for cross-modal retrieval.
- 867, TITLE: Scalable Discriminative Discrete Hashing For Large-Scale Cross-Modal Retrieval  
<https://doi.org/10.1109/ICASSP39728.2021.9413871>  
AUTHORS: J. Qin, L. Fei, J. Zhu, J. Wen, C. Tian and S. Wu  
HIGHLIGHT: In this paper, we propose a supervised discrete-based cross-modal hashing method, named Scalable Discriminative Discrete Hashing (SDDH), for cross-modal retrieval, where 1) the discrete hash codes are directly obtained by multi-modal features and semantic labels so that the quantization errors are dramatically reduced, and 2) the discrete hash codes simultaneously preserve the heterogeneous similarity and manifold information in the original space by employing matrix factoring with orthogonal and balanced constraints.
- 868, TITLE: Hierarchical Similarity Learning for Language-Based Product Image Retrieval  
<https://doi.org/10.1109/ICASSP39728.2021.9415007>  
AUTHORS: Z. Ma, F. Liu, J. Dong, X. Qu, Y. He and S. Ji  
HIGHLIGHT: In this paper, we focus on the cross-modal similarity measurement, and propose a novel Hierarchical Similarity Learning (HSL) network.

- 869, TITLE: Bidirectional Focused Semantic Alignment Attention Network for Cross-Modal Retrieval  
<https://doi.org/10.1109/ICASSP39728.2021.9414382>  
AUTHORS: S. Cheng, L. Wang, A. Du and Y. Li  
HIGHLIGHT: In order to further accurately capture the multi-modal semantic information, a bidirectional focused semantic alignment attention network (BFSAAN) is proposed to handle cross-modal retrieval tasks.
- 870, TITLE: Detection of Audio-Video Synchronization Errors Via Event Detection  
<https://doi.org/10.1109/ICASSP39728.2021.9414924>  
AUTHORS: J. P. Ebeneze, Y. Wu, H. Wei, S. Sethuraman and Z. Liu  
HIGHLIGHT: We present a new method and a large-scale database to detect audio-video synchronization(A/V sync) errors in tennis videos.
- 871, TITLE: FC2RN: A Fully Convolutional Corner Refinement Network for Accurate Multi-Oriented Scene Text Detection  
<https://doi.org/10.1109/ICASSP39728.2021.9413821>  
AUTHORS: X. Qin, Y. Zhou, Y. Guo, D. Wu and W. Wang  
HIGHLIGHT: To address the dilemma, a novel proposed corner-aware convolution in which the sampling positions tightly cover the text area is utilized to encode an initial corner prediction into the feature maps, which can be further used to produce a refined corner prediction.
- 872, TITLE: DoA estimation of a hidden RF source exploiting simple backscatter radio tags  
<https://doi.org/10.1109/ICASSP39728.2021.9414918>  
AUTHORS: G. Vougioukas and A. Bletsas  
HIGHLIGHT: This work emulates a multi-antenna system using a singleantenna receiver and exploiting the beauty and simplicity of backscatter radio.
- 873, TITLE: Probability of Resolution of G-MUSIC: An Asymptotic Approach  
<https://doi.org/10.1109/ICASSP39728.2021.9414025>  
AUTHORS: D. Schenck, X. Mestre and M. Pesavento  
HIGHLIGHT: In this paper, the outlier production mechanism of the G-MUSIC Direction-of-Arrival estimation technique is investigated using tools from Random Matrix Theory.
- 874, TITLE: A Partially-Relaxed Robust DOA Estimator Under Non-Gaussian Low-Rank Interference and Noise  
<https://doi.org/10.1109/ICASSP39728.2021.9413951>  
AUTHORS: M. Trinh-Hoang, M. N. El Korso and M. Pesavento  
HIGHLIGHT: In this paper, we propose a novel robust DOA estimator from the data collected at the sensor array under the corruption of non-Gaussian interference and noise.
- 875, TITLE: Non-Coherent DOA Estimation of Off-Grid Signals With Uniform Circular Arrays  
<https://doi.org/10.1109/ICASSP39728.2021.9415069>  
AUTHORS: Z. Wan and W. Liu  
HIGHLIGHT: However, this may not be correct in practice; in order to address this issue, an off-grid model involved with a bias vector is proposed and an efficient two-step method based on this model is developed.
- 876, TITLE: Enhanced Standard Esprit For Overcoming Imperfections In DOA Estimation  
<https://doi.org/10.1109/ICASSP39728.2021.9413726>  
AUTHORS: M. Esfandiari and S. A. Vorobyov  
HIGHLIGHT: The proposed methods use statistics of the subspace perturbation.
- 877, TITLE: Constrained Tensor Decomposition for 2d DOA Estimation In Transmit Beamspace Mimo Radar with Subarrays  
<https://doi.org/10.1109/ICASSP39728.2021.9413987>  
AUTHORS: F. Xu and S. A. Vorobyov  
HIGHLIGHT: In this paper, a constrained tensor decomposition method that enables two dimensional (2D) direction of arrival (DOA) estimation for transmit beamspace (TB) Multiple-Input Multiple-Output (MIMO) radar with subarrays is proposed.
- 878, TITLE: Alternating Projections Gridless Covariance-Based Estimation For DOA  
<https://doi.org/10.1109/ICASSP39728.2021.9414972>  
AUTHORS: Y. Park and P. Gerstoft

**HIGHLIGHT:** We present a gridless sparse iterative covariance-based estimation method based on alternating projections for direction-of-arrival (DOA) estimation.

879, **TITLE:** Synthetic Data For Dnn-Based Doa Estimation of Indoor Speech

<https://doi.org/10.1109/ICASSP39728.2021.9414415>

**AUTHORS:** F. B. Gelderblom, Y. Liu, J. Kvam and T. A. Myrvoll

**HIGHLIGHT:** This paper investigates the use of different room impulse response (RIR) simulation methods for synthesizing training data for deep neural network-based direction of arrival (DOA) estimation of speech in reverberant rooms. Different sets of synthetic RIRs are obtained using the image source method (ISM) and more advanced methods including diffuse reflections and/or source directivity.

880, **TITLE:** Direction Of Arrival Estimation For Non-Coherent Sub-Arrays Via Joint Sparse And Low-Rank Signal Recovery

<https://doi.org/10.1109/ICASSP39728.2021.9414905>

**AUTHORS:** T. Tirer and O. Bialer

**HIGHLIGHT:** In this paper, we consider a more challenging DOA estimation task where the array is composed of non-coherent sub-arrays (i.e., sub-arrays that observe different unknown phase shifts due to using low-cost unsynchronized local oscillators).

881, **TITLE:** Sparsity And Nonnegativity Constrained Krylov Approach For Direction Of Arrival Estimation

<https://doi.org/10.1109/ICASSP39728.2021.9415040>

**AUTHORS:** H. Baali, A. Bouzerdoun and A. Khelif

**HIGHLIGHT:** This paper presents an approach that imposes two additional constraints to the inverse problem, namely sparsity and nonnegativity of the solution.

882, **TITLE:** Hybrid Analog-Digital MIMO Radar Receivers With Bit-Limited ADCs

<https://doi.org/10.1109/ICASSP39728.2021.9413749>

**AUTHORS:** F. Xi, N. Shlezinger and Y. C. Eldar

**HIGHLIGHT:** In this work we study reduced cost MIMO radar receivers restricted to operate with low resolution ADCs.

883, **TITLE:** Sparse Array Transceiver Design for Enhanced Adaptive Beamforming in MIMO Radar

<https://doi.org/10.1109/ICASSP39728.2021.9414650>

**AUTHORS:** S. A. Hamza, W. Zhai, X. Wang and M. G. Amin

**HIGHLIGHT:** In this paper, we examine the active sparse array design enabling the maximum signal to interference plus noise ratio (MaxSINR) beamforming at the MIMO radar receiver.

884, **TITLE:** Sparse Parameter Estimation for PMCW MIMO Radar Using Few-Bit ADCs

<https://doi.org/10.1109/ICASSP39728.2021.9414267>

**AUTHORS:** C. -Y. Wu, J. Li and T. F. Wong

**HIGHLIGHT:** In this work, we consider target parameter estimation of phase-modulated continuous-wave (PMCW) multiple-input multiple-output (MIMO) radars with few-bit analog-to-digital converters (ADCs).

885, **TITLE:** Parameter Identifiability Of Spatial-Smoothing-Based Bistatic Mimo Radar

<https://doi.org/10.1109/ICASSP39728.2021.9414865>

**AUTHORS:** J. Shi, F. Wen, Y. Liu, Q. Shen, Z. Li and Z. Liu

**HIGHLIGHT:** In this paper, we are devoted to establishing more accurate conditions by studying the positive definiteness of smoothed target covariance matrix.

886, **TITLE:** Parameter Estimation for Coherent Passive MIMO Radar with Unknown Signals under Direct Path Influence

<https://doi.org/10.1109/ICASSP39728.2021.9414746>

**AUTHORS:** Z. Wang and Q. He

**HIGHLIGHT:** This paper studies the problem of joint target position and velocity estimation for a coherent passive radar system.

887, **TITLE:** Riemannian Geometric Optimization Methods for Joint Design of Transmit Sequence and Receive Filter of MIMO Radar

<https://doi.org/10.1109/ICASSP39728.2021.9413665>

**AUTHORS:** J. Li, G. Liao, Y. Huang and A. Nehorai

**HIGHLIGHT:** In this paper, we propose a novel optimization framework to solve the resultant non-convex problem on a Riemannian product manifold.

- 888, TITLE: High Accuracy Tracking of Targets Using Massive MIMO  
<https://doi.org/10.1109/ICASSP39728.2021.9414089>  
AUTHORS: X. Zeng, F. Zhang, B. Wang and K. J. Ray Liu  
HIGHLIGHT: This paper proposes an accurate and novel multipath-resilient system to track the targets by leveraging the large number of antennas in massive MIMO systems.
- 889, TITLE: Admm-Based Fast Algorithm for Robust Multi-Group Multicast Beamforming  
<https://doi.org/10.1109/ICASSP39728.2021.9413651>  
AUTHORS: N. Mohamadi, M. Dong and S. ShahbazPanahi  
HIGHLIGHT: We develop an alternating direction method of multipliers (ADMM)based fast algorithm to solve this problem directly with convergence guarantee.
- 890, TITLE: Scalable and Distributed MMSE Algorithms for Uplink Receive Combining in Cell-Free Massive MIMO Systems  
<https://doi.org/10.1109/ICASSP39728.2021.9414456>  
AUTHORS: R. Van Rompaey and M. Moonen  
HIGHLIGHT: In this paper, the problem of optimal uplink receive combining is tackled by providing an efficient distributed MMSE algorithm, with a minimal number of exchanged parameters between the APs and the network center.
- 891, TITLE: Antenna Selection for Massive MIMO Systems Based on POMDP Framework  
<https://doi.org/10.1109/ICASSP39728.2021.9414692>  
AUTHORS: S. Sharifi, S. Shahbaz Panahi and M. Dong  
HIGHLIGHT: To avoid the computational complexity associated with the value iteration algorithm, we herein propose to use the simple myopic antenna selection policy based on the fact that for any arbitrary number of antennas and RF chains, under the assumption of positively correlated two-state Markov channel model, the myopic policy is optimal.
- 892, TITLE: RIS-Aided Joint Localization and Synchronization with a Single-Antenna Mmwave Receiver  
<https://doi.org/10.1109/ICASSP39728.2021.9413515>  
AUTHORS: A. Fascista, A. Coluccia, H. Wymeersch and G. Seco-Granados  
HIGHLIGHT: In this paper we show that thanks to the use of a reconfigurable intelligent surface (RIS), joint localization and synchronization is possible with only downlink MISO transmissions.
- 893, TITLE: Joint Channel, Data, and Phase-Noise Estimation in MIMO-OFDM Systems Using a Tensor Modeling Approach  
<https://doi.org/10.1109/ICASSP39728.2021.9414541>  
AUTHORS: B. Sokal, P. R. B. Gomes, A. L. F. de Almeida and M. Haardt  
HIGHLIGHT: In this work, we propose a two-stage tensor-based receiver for joint channel, phase-noise (PN), and data estimation in MIMO-OFDM systems.
- 894, TITLE: Robust Steerable Differential Beamformers with Null Constraints for Concentric Circular Microphone Arrays  
<https://doi.org/10.1109/ICASSP39728.2021.9414119>  
AUTHORS: X. Wang, G. Huang, I. Cohen, J. Benesty and J. Chen  
HIGHLIGHT: In this paper, we extend our recently developed beamforming method for CMAs to the design of differential beamformers with CCMAAs, which takes advantage of the symmetric null constraints from the beampattern.
- 895, TITLE: Close-Talking Recording with Planarly Distributed Microphones  
<https://doi.org/10.1109/ICASSP39728.2021.9414828>  
AUTHORS: T. Okamoto  
HIGHLIGHT: This paper provides a close-talking recording method with microphones in arbitrary planar distributions based on sound pressure interpolation.
- 896, TITLE: (W)earable Microphone Array and Ultrasonic Echo Localization for Coarse Indoor Environment Mapping  
<https://doi.org/10.1109/ICASSP39728.2021.9414356>  
AUTHORS: F. Pfeundtner, J. Yang and G. S?r?s  
HIGHLIGHT: We present a microphone array structure for spherical sound incidence angle tracking that can be attached to headphones or directly integrated into earphones.
- 897, TITLE: Characterization of Mems Microphone Sensitivity and Phase Distributions with Applications in Array Processing  
<https://doi.org/10.1109/ICASSP39728.2021.9413352>

- AUTHORS: P. W. A. Wijnings, S. Stuijk, R. Scholte and H. Corporaal  
HIGHLIGHT: Hence, this work demonstrates a free-field comparison method for measuring these variations in a batch of arrays.
- 898, TITLE: Directional Sparse Filtering Using Weighted Lehmer Mean for Blind Separation of Unbalanced Speech Mixtures  
<https://doi.org/10.1109/ICASSP39728.2021.9414336>  
AUTHORS: K. Watcharasupat, A. H. T. Nguyen, C. -H. Ooi and A. W. H. Khong  
HIGHLIGHT: We propose an algorithm based on the directional sparse filtering (DSF) framework that utilizes the Lehmer mean with learnable weights to adaptively account for source imbalance.
- 899, TITLE: Distributed Speech Separation in Spatially Unconstrained Microphone Arrays  
<https://doi.org/10.1109/ICASSP39728.2021.9414758>  
AUTHORS: N. Furnon, R. Serizel, I. Illina and S. Essid  
HIGHLIGHT: We propose a distributed algorithm that can process spatial information in a spatially unconstrained microphone array.
- 900, TITLE: An Adaptive Non-Linear Process for Under-Determined Virtual Microphone Beamforming  
<https://doi.org/10.1109/ICASSP39728.2021.9413813>  
AUTHORS: M. Bekrani, A. H. T. Nguyen and A. W. H. Khong  
HIGHLIGHT: We propose to extend existing virtual microphone signal interpolation by employing an adaptive non-linear (ANL) process for acoustic beamforming.
- 901, TITLE: Window Beamformer for Sparse Concentric Circular Array  
<https://doi.org/10.1109/ICASSP39728.2021.9414069>  
AUTHORS: R. Sharma, I. Cohen and B. Berdugo  
HIGHLIGHT: This work proposes a practical fixed beamformer for a sparse concentric circular array (CCA) of microphones.
- 902, TITLE: Single-Point Array Response Control with Minimum Pattern Deviation  
<https://doi.org/10.1109/ICASSP39728.2021.9413763>  
AUTHORS: X. Ai and L. Gan  
HIGHLIGHT: This paper presents a beam pattern synthesis scheme based on the single-point array response control with minimum pattern deviation (SPARC-MPD) method.
- 903, TITLE: Focusing-Based Wideband Adaptive Beamforming Using Covariance Matrix Reconstruction  
<https://doi.org/10.1109/ICASSP39728.2021.9414963>  
AUTHORS: P. Chen, W. Wang and J. Gao  
HIGHLIGHT: To maximize the overall performance of focusing-based beamformer, this paper presents an adaptive focusing transformation-based beamforming algorithm using covariance matrix reconstruction.
- 904, TITLE: Bayesian Multiple Change-Point Detection of Propagating Events  
<https://doi.org/10.1109/ICASSP39728.2021.9413434>  
AUTHORS: T. Halme, E. Nitzan and V. Koivunen  
HIGHLIGHT: In this paper, we take a Bayesian approach and model the detection of spatial events as a Bayesian multiple change point detection problem.
- 905, TITLE: One-Bit Autocorrelation Estimation With Non-Zero Thresholds  
<https://doi.org/10.1109/ICASSP39728.2021.9414732>  
AUTHORS: C. -L. Liu and Z. -M. Lin  
HIGHLIGHT: This paper presents an autocorrelation estimator based on a one-bit quantizer with a non-zero threshold.
- 906, TITLE: A Novel Bayesian Approach for the Two-Dimensional Harmonic Retrieval Problem  
<https://doi.org/10.1109/ICASSP39728.2021.9414486>  
AUTHORS: R. R. Pote and B. D. Rao  
HIGHLIGHT: In this work we propose a novel Bayesian strategy to address the two dimensional harmonic retrieval problem, through remodeling and reparameterization of the standard data model.
- 907, TITLE: On Overfitting in Discrete Super-Resolution Recovery  
<https://doi.org/10.1109/ICASSP39728.2021.9415033>  
AUTHORS: W. Lu and H. Qiao



HIGHLIGHT: This paper studies the overfitting in discrete super-resolution problem.

908, TITLE: SIML: Sieved Maximum Likelihood for Array Signal Processing

<https://doi.org/10.1109/ICASSP39728.2021.9414991>

AUTHORS: M. Simeoni and P. Hurley

HIGHLIGHT: In this work, we propose a Sieved Maximum Likelihood (SiML) method.

909, TITLE: Estimation of Groundwater Storage Variations in Indus River Basin Using Grace Data

<https://doi.org/10.1109/ICASSP39728.2021.9413591>

AUTHORS: Y. Sattar and Z. Khalid

HIGHLIGHT: We present a data processing framework that processes and combines these data-sets to provide an estimate of GWS changes.

910, TITLE: Temporal Exemplar Channels In High-Multipath Environments

<https://doi.org/10.1109/ICASSP39728.2021.9414262>

AUTHORS: M. Kashef, P. Vouras, R. Jones, R. Candell and K. A. Remley

HIGHLIGHT: As a result, in this work, we introduce a machine learning (ML) based exemplar extraction approach on mmWave wireless spatial-channel measurements.

911, TITLE: Multi-Vehicle Velocity Estimation Using IEEE 802.11ad Waveform

<https://doi.org/10.1109/ICASSP39728.2021.9413567>

AUTHORS: G. Han, S. Kim and J. Choi

HIGHLIGHT: In this paper, we propose a multi-target velocity estimation technique using IEEE 802.11ad waveform in a vehicle-to-vehicle (V2V) scenario.

912, TITLE: Real-Time Interaural Time Delay Estimation via Onset Detection

<https://doi.org/10.1109/ICASSP39728.2021.9414632>

AUTHORS: E. Ren, G. C. Ornelas and H. -A. Loeliger

HIGHLIGHT: This paper introduces a novel real-time estimation method inspired by the precedence effect.

913, TITLE: EKNet: Learning System Noise Statistics from Measurement Data

<https://doi.org/10.1109/ICASSP39728.2021.9415083>

AUTHORS: L. Xu and R. Niu

HIGHLIGHT: In this paper, to reduce the time and manpower spent on fine-tuning an extended Kalman filter (EKF), we propose a new learning framework, EKNet, for automatically estimating the best process and measurement noise covariance pair from the real measurement data.

914, TITLE: Sliding-Capon Based Convolutional BeamSpace for Linear Arrays

<https://doi.org/10.1109/ICASSP39728.2021.9414022>

AUTHORS: P. -C. Chen and P. P. Vaidyanathan

HIGHLIGHT: A new method to design the filter for convolutional beamspace (CBS), called Capon-CBS, is proposed.

915, TITLE: Target Detection from Distributed Passive Sensors: Semi-Labeled Data Quantization

<https://doi.org/10.1109/ICASSP39728.2021.9413917>

AUTHORS: Z. Sutton, P. Willett and S. Marano

HIGHLIGHT: This work considers the case where both the intensity data and the label (distance) values are coarsely quantized to decrease communication cost.

916, TITLE: Sparse Factorization-Based Detection of Off-the-Grid Moving Targets Using FMCW Radars

<https://doi.org/10.1109/ICASSP39728.2021.9414925>

AUTHORS: G. M. de Galland, T. Feuillen, L. Vandendorpe and L. Jacques

HIGHLIGHT: In this paper, we investigate the application of continuous sparse signal reconstruction algorithms for the estimation of the ranges and speeds of multiple moving targets using an FMCW radar.

917, TITLE: A Robust Copula Model for Radar-Based Landmine Detection

<https://doi.org/10.1109/ICASSP39728.2021.9413694>

AUTHORS: A. D. Pambudi, F. Ahmad and A. M. Zoubir

HIGHLIGHT: We present a robust copula model for landmine detection based on a likelihood ratio test.

- 918, TITLE: Radar Clutter Classification Using Expectation-Maximization Method  
<https://doi.org/10.1109/ICASSP39728.2021.9413918>  
AUTHORS: S. Han, P. Addabbo, D. Orlando and G. Ricci  
HIGHLIGHT: In this paper, the problem of classifying radar clutter returns into statistically homogeneous subsets is addressed.
- 919, TITLE: A Meta-Learning Framework for Few-Shot Classification of Remote Sensing Scene  
<https://doi.org/10.1109/ICASSP39728.2021.9413971>  
AUTHORS: P. Zhang, Y. Bai, D. Wang, B. Bai and Y. Li  
HIGHLIGHT: In this work, we provide a metalearning framework for few-shot classification of RS scene.
- 920, TITLE: Differential Convolution Feature Guided Deep Multi-Scale Multiple Instance Learning for Aerial Scene Classification  
<https://doi.org/10.1109/ICASSP39728.2021.9414323>  
AUTHORS: B. Zhou, J. Yi and Q. Bi  
HIGHLIGHT: In this paper, we propose a deep multi-scale multiple instance learning (DMSMIL) framework to tackle the above challenges.
- 921, TITLE: Generalized Thinned Coprime Array for DOA Estimation  
<https://doi.org/10.1109/ICASSP39728.2021.9414146>  
AUTHORS: J. Shi, Y. Liu, F. Wen, Z. Liu, P. Hu and Z. Gong  
HIGHLIGHT: In this paper, we propose a generalized thinned coprime array by introducing the flexible inter-element spacings, where the conventional one can be seen as a special case.
- 922, TITLE: TCLA Array: A New Sparse Array Design with Less Mutual Coupling  
<https://doi.org/10.1109/ICASSP39728.2021.9415108>  
AUTHORS: A. M. A. Shaalan, J. Du and Y. -H. Tu  
HIGHLIGHT: This paper aims to introduce a new proposed sparse array that takes into account all these considerations.
- 923, TITLE: Low Mutual Coupling Sparse Array Design Using ULA Fitting  
<https://doi.org/10.1109/ICASSP39728.2021.9414768>  
AUTHORS: W. Shi, Y. Li and S. A. Vorobyov  
HIGHLIGHT: In this paper, a general sparse array (SA) design principle, called uniform linear array (ULA) fitting, is proposed.
- 924, TITLE: Low-Rank and Sparse Decomposition for Joint DOA Estimation and Contaminated Sensors Detection with Sparsely Contaminated Arrays  
<https://doi.org/10.1109/ICASSP39728.2021.9413843>  
AUTHORS: H. Huang and A. M. Zoubir  
HIGHLIGHT: In this work, we consider sparsely contaminated arrays in which only a few sensors are contaminated by sensor gain and phase errors, and moreover, the number of contaminated sensors as well as their positions are unknown.
- 925, TITLE: Fundamental Trade-Offs in Noisy Super-Resolution with Synthetic Apertures  
<https://doi.org/10.1109/ICASSP39728.2021.9413426>  
AUTHORS: S. Shahsavari, J. Millhiser and P. Pal  
HIGHLIGHT: The main contribution of this paper is to rigorously establish that nested arrays provide lower Cramér-Rao bounds than a ULA (with the same number of sensors) in the low SNR regime, and therefore can lead to better resolvability of closely spaced sources.
- 926, TITLE: Enhanced Blind Calibration of Uniform Linear Arrays with One-Bit Quantization by Kullback-Leibler Divergence Covariance Fitting  
<https://doi.org/10.1109/ICASSP39728.2021.9413647>  
AUTHORS: A. Weiss and A. Yeredor  
HIGHLIGHT: In this paper, we present a novel solution approach for the blind calibration problem, namely, without using known calibration signals.
- 927, TITLE: Non-Iterative Blind Calibration of Nested Arrays with Asymptotically Optimal Weighting  
<https://doi.org/10.1109/ICASSP39728.2021.9415037>  
AUTHORS: A. Weiss and A. Yeredor  
HIGHLIGHT: In this paper, we present a novel blind calibration method for 2-level nested arrays.

- 928, TITLE: Sensor Networks TDOA Self-Calibration: 2D Complexity Analysis and Solutions  
<https://doi.org/10.1109/ICASSP39728.2021.9414634>  
AUTHORS: L. Ferranti, K. ?str?m, M. Oskarsson, J. Boutellier and J. Kannala  
HIGHLIGHT: In this paper we consider 2D networks with synchronized receivers but unsynchronized transmitters and the corresponding calibration techniques, known as Time-Difference-Of-Arrival (TDOA) techniques.
- 929, TITLE: Fast and Robust Stratified Self-Calibration Using Time-Difference-Of-Arrival Measurements  
<https://doi.org/10.1109/ICASSP39728.2021.9414309>  
AUTHORS: M. Larsson, G. Flood, M. Oskarsson and K. ?str?m  
HIGHLIGHT: In this paper we study the problem of estimating receiver and sender positions using time-difference-of-arrival measurements.
- 930, TITLE: Stability Analysis of the RC-PLMS Adaptive Beamformer Using a Simple Transfer Function Approximation  
<https://doi.org/10.1109/ICASSP39728.2021.9413445>  
AUTHORS: G. Akkad, A. Mansour, B. El Hassan and E. Inaty  
HIGHLIGHT: In this paper, we propose a discrete time transfer function approximation for the reduced complexity parallel least mean square (RC-pLMS) adaptive beamforming algorithm.
- 931, TITLE: On The Asymptotic Performance of One-Bit Co-Array-Based Music  
<https://doi.org/10.1109/ICASSP39728.2021.9414706>  
AUTHORS: S. Sedighi, B. Shankar, M. Soltanalian and B. Ottersten  
HIGHLIGHT: In this paper, to provide valuable insights into the performance of DoA estimation from one-bit SLA measurements, we derive an asymptotic closed-form expression for the performance of One-Bit Co-Array-Based MUSIC (OBCAB-MUSIC).
- 932, TITLE: KId Minimization-Based Constrained Measurement Filtering For Two-Step TDOA Indoor Tracking  
<https://doi.org/10.1109/ICASSP39728.2021.9413374>  
AUTHORS: R. Huang, L. Yang, J. Tao and Y. Xue  
HIGHLIGHT: This paper presents an enhanced two-step method for tracking an indoor point target using the time difference of arrival (TDOA) measurements from an ultra wideband (UWB) positioning system.
- 933, TITLE: A Correntropy Based Algorithm for Robust Localization in Wireless Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414143>  
AUTHORS: M. Sedighizad, B. Seyfe and S. Valaee  
HIGHLIGHT: In this paper, a correntropy based algorithm is proposed for localization and tracking of a mobile station in wireless networks.
- 934, TITLE: MuG: A Multipath-Exploited and Grid-Free Localisation Method  
<https://doi.org/10.1109/ICASSP39728.2021.9414329>  
AUTHORS: H. Liu, W. Dai and Y. Shen  
HIGHLIGHT: This paper presents a localisation method in which both LoS and NLoS paths are exploited for much more general settings.
- 935, TITLE: Sparse Bayesian Learning for Acoustic Source Localization  
<https://doi.org/10.1109/ICASSP39728.2021.9413960>  
AUTHORS: R. Pandey, S. Nannuru and A. Siripuram  
HIGHLIGHT: In this paper, the CS method of sparse Bayesian learning (SBL) is used to find the DOAs.
- 936, TITLE: Automatic Fine-Grained Localization of Utility Pole Landmarks on Distributed Acoustic Sensing Traces Based on Bilinear Resnets  
<https://doi.org/10.1109/ICASSP39728.2021.9415049>  
AUTHORS: Y. Lu, Y. Tian, S. Han, E. Cosatto, S. Ozharar and Y. Ding  
HIGHLIGHT: In this paper, we propose two machine learning approaches to automate this procedure for large-scale implementation.
- 937, TITLE: SSLIDE: Sound Source Localization for Indoors Based on Deep Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9415109>  
AUTHORS: Y. Wu, R. Ayyalasomayajula, M. J. Bianco, D. Bharadia and P. Gerstoft  
HIGHLIGHT: This paper presents SSLIDE, Sound Source Localization for Indoors using DEep learning, which applies deep neural networks (DNNs) with encoder-decoder structure to localize sound sources with random positions in a continuous space.

938, TITLE: Physical-Layer Security via Distributed Beamforming in the Presence of Adversaries with Unknown Locations  
<https://doi.org/10.1109/ICASSP39728.2021.9414861>  
AUTHORS: Y. Savas, A. Hashemi, A. P. Vinod, B. M. Sadler and U. Topcu  
HIGHLIGHT: We study the problem of securely communicating a sequence of information bits with a client in the presence of multiple adversaries at unknown locations in the environment.

939, TITLE: Canonical Polyadic Tensor Decomposition With Low-Rank Factor Matrices  
<https://doi.org/10.1109/ICASSP39728.2021.9414606>  
AUTHORS: A. -H. Phan, P. Tichavsk?, K. Sobolev, K. Sozykin, D. Ermilov and A. Cichocki  
HIGHLIGHT: This paper proposes a constrained canonical polyadic (CP) tensor decomposition method with low-rank factor matrices.

940, TITLE: A Diffusion FXLMS Algorithm for Multi-Channel Active Noise Control and Variable Spatial Smoothing  
<https://doi.org/10.1109/ICASSP39728.2021.9414609>  
AUTHORS: Y. J. Chu, S. C. Chan, C. M. Mak and M. Wu  
HIGHLIGHT: To solve this problem, a new Diff filtered-x least mean squares (Diff-FxLMS) algorithm that incorporates the knowledge of spatial smoothness is proposed.

941, TITLE: ADAPT-Then-Combine Full Waveform Inversion for Distributed Subsurface Imaging In Seismic Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414072>  
AUTHORS: B. -S. Shin and D. Shutin  
HIGHLIGHT: To this end, we propose a distributed implementation of the full waveform inversion (FWI) for distributed imaging of subsurfaces in seismic networks.

942, TITLE: Data Fusion for Audiovisual Speaker Localization: Extending Dynamic Stream Weights to the Spatial Domain  
<https://doi.org/10.1109/ICASSP39728.2021.9413399>  
AUTHORS: J. Wissing et al.  
HIGHLIGHT: This paper proposes a novel audiovisual data fusion framework for speaker localization by assigning individual dynamic stream weights to specific regions in the localization space.

943, TITLE: Towards Robust Training of Multi-Sensor Data Fusion Network Against Adversarial Examples in Semantic Segmentation  
<https://doi.org/10.1109/ICASSP39728.2021.9413772>  
AUTHORS: Y. Yu, H. J. Lee, B. C. Kim, J. U. Kim and Y. M. Ro  
HIGHLIGHT: To achieve adversarial robust multi-sensor data fusion networks, we propose here a novel robust training scheme called Multi-Sensor Cumulative Learning (MSCL).

944, TITLE: Data-Driven Adaptive Network Resource Slicing for Multi-Tenant Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9413706>  
AUTHORS: N. Reyhanian, H. Farmanbar and Z. -Q. Luo  
HIGHLIGHT: In this paper, we propose a novel framework for network slicing with the goal of maximizing the expected utilities of tenants in the backhaul and Radio Access Network (RAN), where we reconfigure slices according to the time-varying user traffic and channel states.

945, TITLE: Distributed Scheduling Using Graph Neural Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414098>  
AUTHORS: Z. Zhao, G. Verma, C. Rao, A. Swami and S. Segarra  
HIGHLIGHT: To overcome this limitation, we propose a distributed MWIS solver based on graph convolutional networks (GCNs).

946, TITLE: Efficient Power Allocation Using Graph Neural Networks and Deep Algorithm Unfolding  
<https://doi.org/10.1109/ICASSP39728.2021.9415106>  
AUTHORS: A. Chowdhury, G. Verma, C. Rao, A. Swami and S. Segarra  
HIGHLIGHT: We study the problem of optimal power allocation in a single-hop ad hoc wireless network.

947, TITLE: A Sample-Efficient Scheme for Channel Resource Allocation in Networked Estimation  
<https://doi.org/10.1109/ICASSP39728.2021.9414552>  
AUTHORS: M. M. Vasconcelos and U. Mitra

**HIGHLIGHT:** Leveraging symmetry, quasi-convexity, and the method of Kernel density estimation, we propose a data-driven algorithm, which is guaranteed to converge to a globally optimal solution.

948, **TITLE:** An Efficient Linear Programming Rounding-and-Refinement Algorithm for Large-Scale Network Slicing Problem

<https://doi.org/10.1109/ICASSP39728.2021.9413378>

**AUTHORS:** W. -K. Chen, Y. -F. Liu, Y. -H. Dai and Z. -Q. Luo

**HIGHLIGHT:** In this paper, we consider the network slicing problem which attempts to map multiple customized virtual network requests (also called services) to a common shared network infrastructure and allocate network resources to meet diverse service requirements, and propose an efficient two-stage algorithm for solving this NP-hard problem.

949, **TITLE:** Efficient Migration to the Next Generation of Networks Based on Digital Annealing

<https://doi.org/10.1109/ICASSP39728.2021.9414531>

**AUTHORS:** M. Javad-Kalbasi and S. Valaee

**HIGHLIGHT:** In this paper, the network migration problem is considered as a set of circuit migration problems in which two technicians simultaneously migrate the two ends of a circuit in order to minimize the total accumulated sites in-service and total technician travels.

950, **TITLE:** A Technique for OFDM Symbol Slicing

<https://doi.org/10.1109/ICASSP39728.2021.9414504>

**AUTHORS:** A. Perez-Neira and M. A. Lagunas

**HIGHLIGHT:** This work presents an orthonormal transform that splits the Orthogonal Frequency Division Multiplex (OFDM) symbol into slices with ranked rate and decoding complexity.

951, **TITLE:** Communication Over Block Fading Channels ? An Algorithmic Perspective On Optimal Transmission Schemes

<https://doi.org/10.1109/ICASSP39728.2021.9413887>

**AUTHORS:** H. Boche, R. F. Schaefer and H. Vincent Poor

**HIGHLIGHT:** This paper approaches this issue from a fundamental, algorithmic point of view by studying whether or not it is in principle possible to construct or find such optimal transmission schemes algorithmically (without putting any constraints on the computational complexity of such algorithms).

952, **TITLE:** Secure UAV Communications Under Uncertain Eavesdroppers Locations

<https://doi.org/10.1109/ICASSP39728.2021.9413480>

**AUTHORS:** S. Wang, F. Kong and Q. Li

**HIGHLIGHT:** In this paper, we consider the UAV-enabled physical-layer secure communications.

953, **TITLE:** On Strategic Jamming in Distributed Detection Networks

<https://doi.org/10.1109/ICASSP39728.2021.9414726>

**AUTHORS:** C. Quan, B. Geng and P. K. Varshney

**HIGHLIGHT:** In this paper, the optimal jamming strategy by an adversary in distributed detection networks is investigated.

954, **TITLE:** Real Number Signal Processing can Detect Denial-of-Service Attacks

<https://doi.org/10.1109/ICASSP39728.2021.9413911>

**AUTHORS:** H. Boche, R. F. Schaefer and H. Vincent Poor

**HIGHLIGHT:** This paper investigates the general computation framework of Blum-Shub-Smale machines which allows the processing and storage of arbitrary reals.

955, **TITLE:** A Hybrid Approach to Coded Compressed Sensing Where Coupling Takes Place Via the Outer Code

<https://doi.org/10.1109/ICASSP39728.2021.9414469>

**AUTHORS:** J. R. Ebert, V. K. Amalladinne, J. -F. Chamberland and K. R. Narayanan

**HIGHLIGHT:** This article seeks to advance coded compressed sensing (CCS) as a practical scheme for unsourced random access.

956, **TITLE:** Globally Optimal Beamforming for Rate Splitting Multiple Access

<https://doi.org/10.1109/ICASSP39728.2021.9413616>

**AUTHORS:** B. Matthiesen, Y. Mao, P. Popovski and B. Clerckx

**HIGHLIGHT:** We consider globally optimal precoder design for rate splitting multiple access in Gaussian multiple-input single-output downlink channels with respect to weighted sum rate and energy efficiency maximization.

957, **TITLE:** Beam Focusing for Multi-User MIMO Communications with Dynamic Metasurface Antennas

<https://doi.org/10.1109/ICASSP39728.2021.9413746>

AUTHORS: H. Zhang, N. Shlezinger, F. Guidi, D. Dardari, M. F. Imani and Y. C. Eldar

HIGHLIGHT: In this paper, we study the potential of beam focusing, feasible in near-field operation, for multi-user MIMO systems, where the base station is equipped with a DMA.

958, TITLE: Pushing The Limit of Type I Codebook For Fdd Massive MIMO Beamforming: A Channel Covariance Reconstruction Approach

<https://doi.org/10.1109/ICASSP39728.2021.9414376>

AUTHORS: K. Li, Y. Li, L. Cheng, Q. Shi and Z. -Q. Luo

HIGHLIGHT: Therefore, in this paper, using Type I codebook, we leverage advances in cutting plane method to optimize the CSI reconstruction at the base station (BS), in order to close the gap between these two codebook based beamforming schemes.

959, TITLE: First-Order Fast Algorithm for Structurally Optimal Multi-Group Multicast Beamforming in Large-Scale Systems

<https://doi.org/10.1109/ICASSP39728.2021.9414396>

AUTHORS: C. Zhang, M. Dong and B. Liang

HIGHLIGHT: Based on the optimal multicast beamforming structure, we propose a fast first-order algorithm to obtain the beamforming solution.

960, TITLE: Analog Beamforming With Antenna Selection For Large-Scale Antenna Arrays

<https://doi.org/10.1109/ICASSP39728.2021.9414673>

AUTHORS: A. Arora, C. G. Tsinos, B. Shankar Mysore R, S. Chatzinotas and B. Ottersten

HIGHLIGHT: In this work, we propose a joint design of analog beamforming with antenna selection (AS) or antenna placement (AP) for an analog beamforming system.

961, TITLE: Beamforming for Bidirectional MIMO Full Duplex Under the Joint Sum Power and Per Antenna Power Constraints

<https://doi.org/10.1109/ICASSP39728.2021.9414995>

AUTHORS: C. K. Sheemar and D. Slock

HIGHLIGHT: We propose a novel beamforming design to maximize the weighted sum-rate (WSR) with alternating optimization under the limited dynamic range (LDR) noise model.

962, TITLE: Iterative Reweighted Algorithms for Joint User Identification and Channel Estimation in Spatially Correlated Massive MTC

<https://doi.org/10.1109/ICASSP39728.2021.9413733>

AUTHORS: H. Djelouat, M. Leinonen and M. Juntti

HIGHLIGHT: We develop a computationally efficient alternating direction method of multipliers (ADMM) approach to solve it.

963, TITLE: Millimeter Wave MIMO Channel Estimation with 1-bit Spatial Sigma-Delta Analog-to-Digital Converters

<https://doi.org/10.1109/ICASSP39728.2021.9414010>

AUTHORS: R. S. P. Sankar and S. Prabhakar Chepuri

HIGHLIGHT: Therefore, we present a new method for modeling the quantization noise by leveraging the deterministic input-output relation of the 1-bit spatial sigma-delta modulator.

964, TITLE: An Efficient Algorithm For Device Detection And Channel Estimation In Asynchronous IOT Systems

<https://doi.org/10.1109/ICASSP39728.2021.9413870>

AUTHORS: L. Liu and Y. -F. Liu

HIGHLIGHT: This paper targets at two practical issues along this line that have not been addressed before: asynchronous transmission from uncoordinated users and efficient algorithms for real-time implementation in systems with a massive number of devices.

965, TITLE: Kalman Filter Based MIMO CSI Phase Recovery for COTS Wifi Devices

<https://doi.org/10.1109/ICASSP39728.2021.9413408>

AUTHORS: C. Li, J. Brauer, A. Sezgin and C. Zenger

HIGHLIGHT: In this paper, we directly utilize the modeling of the phase distortions caused by the hardware impairments and propose an adaptive CSI estimation approach based on Kalman filter (KF) with maximum-a-posteriori (MAP) estimation that considers the CSI from the previous time.

966, TITLE: Improved Atomic Norm Based Channel Estimation for Time-Varying Narrowband Leaked Channels

<https://doi.org/10.1109/ICASSP39728.2021.9413804>

AUTHORS: J. Li and U. Mitra  
HIGHLIGHT: In this paper, improved channel gain delay estimation strategies are investigated when practical pulse shapes with finite block length and transmission bandwidth are employed.

967, TITLE: Bayesian Massive MIMO Channel Estimation with Parameter Estimation Using Low-Resolution ADCs  
<https://doi.org/10.1109/ICASSP39728.2021.9413477>  
AUTHORS: S. Huang, D. Qiu and T. D. Tran  
HIGHLIGHT: In this paper, we treat both signals and parameters as random variables and recover them jointly within the AMP framework.

968, TITLE: Optimal Detection in the Presence of Non-Gaussian Jamming  
<https://doi.org/10.1109/ICASSP39728.2021.9414769>  
AUTHORS: K. A. Almahorg and R. H. Gohary  
HIGHLIGHT: We consider a scenario in which a transmitter sends complex multidimensional symbols to a receiver in the presence of a proactive continuous jammer emitting a zero-mean complex Gaussian signal over an unknown complex Gaussian channel.

969, TITLE: An Efficient Active Set Algorithm for Covariance Based Joint Data and Activity Detection for Massive Random Access with Massive MIMO  
<https://doi.org/10.1109/ICASSP39728.2021.9413525>  
AUTHORS: Z. Wang, Z. Chen, Y. -F. Liu, F. Sohrab and W. Yu  
HIGHLIGHT: This paper proposes a computationally efficient algorithm to solve the joint data and activity detection problem for massive random access with massive multiple-input multiple-output (MIMO).

970, TITLE: Neural Layered Min-Sum Decoding for Protograph LDPC Codes  
<https://doi.org/10.1109/ICASSP39728.2021.9414543>  
AUTHORS: D. Zhang et al.  
HIGHLIGHT: In this paper, layered min-sum (MS) iterative decoding is formulated as a customized neural network following the sequential scheduling of check node (CN) updates.

971, TITLE: Integer Carrier Frequency Offset Estimation in OFDM with Zadoff-Chu Sequences  
<https://doi.org/10.1109/ICASSP39728.2021.9413937>  
AUTHORS: J. D. Roth, D. A. Garren and R. C. Robertson  
HIGHLIGHT: To address this, we present a Zadoff-Chu sequence-based approach to overcome the ambiguity associated with estimating very high levels of integer Doppler.

972, TITLE: Plug-And-Play Learned Gaussian-mixture Approximate Message Passing  
<https://doi.org/10.1109/ICASSP39728.2021.9414910>  
AUTHORS: O. Musa, P. Jung and G. Caire  
HIGHLIGHT: In this paper, we propose learned Gaussian-mixture AMP (L-GM-AMP) - a plug-and-play compressed sensing (CS) recovery algorithm suitable for any i.i.d. source prior.

973, TITLE: Low-Latency Polar Decoder Using Overlapped SCL Processing  
<https://doi.org/10.1109/ICASSP39728.2021.9414326>  
AUTHORS: D. Kam, B. Y. Kong and Y. Lee  
HIGHLIGHT: In this paper, we present a novel scheduling method that reduces the latency of polar decoders significantly.

974, TITLE: Modular Binary Tree Architecture for Distributed Large Intelligent Surface  
<https://doi.org/10.1109/ICASSP39728.2021.9414989>  
AUTHORS: J. V. Alegría, F. Rusek, J. R. S?nchez and O. Edfors  
HIGHLIGHT: We present a modular architecture that allows combining different LIS panels using a binary tree.

975, TITLE: Stochastic Successive Weighted Sum-Rate Maximization for Multiuser MIMO Systems with Finite-Alphabet Inputs  
<https://doi.org/10.1109/ICASSP39728.2021.9413799>  
AUTHORS: X. Guan, X. Zhao and Q. Shi  
HIGHLIGHT: Inspired by the stochastic successive upper-bound minimization (SSUM) method [1], this paper proposes a stochastic successive inexact lower-bound maximization (SSILM) algorithm for the WSRM problem with finite-alphabet inputs.

976, TITLE: Rate 1 Quasi Orthogonal Universal Transmission and Combining for MIMO Systems Achieving Full Diversity



<https://doi.org/10.1109/ICASSP39728.2021.9414982>

AUTHORS: B. Avraham, U. Erez and E. Domanovitz

HIGHLIGHT: This work addresses general multiple-input multiple-output systems and develops combined diversity transmission and combining schemes that achieve rate one and full diversity with reduced decoding complexity, while being universal in the sense that the operations performed at both transmission ends are channel independent.

977, TITLE: Energy Efficiency Optimization Technique for SWIPT-Enabled Multi-Group Multicasting Systems with Heterogeneous Users

<https://doi.org/10.1109/ICASSP39728.2021.9415060>

AUTHORS: S. Gautam, S. Chatzinotas and B. Ottersten

HIGHLIGHT: An algorithm based on Dinkelback method, slack-variable replacement, and second-order conic programming (SOCP)/semi-definite relaxation (SDR) is proposed to obtain a suitable solution for the above-mentioned fractional-objective dependent non-convex problem.

978, TITLE: Multi-Branch Tomlinson-Harashima Precoding for Rate Splitting Based Systems with Multiple Antennas

<https://doi.org/10.1109/ICASSP39728.2021.9413886>

AUTHORS: A. R. Flores, R. C. de Lamare and B. Clerckx

HIGHLIGHT: In this work, we propose a multi-branch (MB) scheme for a RS-based multiple-antenna system, which creates patterns to order the transmitted symbols and enhances the overall sum rate performance compared to existing approaches.

979, TITLE: Divide and Conquer: One-bit MIMO-OFDM Detection by Inexact Expectation Maximization

<https://doi.org/10.1109/ICASSP39728.2021.9414559>

AUTHORS: M. Shao and W. -K. Ma

HIGHLIGHT: In this paper we study the expectation-maximization (EM) approach for one-bit MIMO-OFDM detection.

980, TITLE: Differential Chaos Shift Keying-Based Wireless Power Transfer

<https://doi.org/10.1109/ICASSP39728.2021.9414014>

AUTHORS: P. Mukherjee, C. Psomas and I. Krikidis

HIGHLIGHT: In this work, we investigate differential chaos shift keying (DCSK), a communication-based waveform, in the context of wireless power transfer (WPT).

981, TITLE: VGAI: End-to-End Learning of Vision-Based Decentralized Controllers for Robot Swarms

<https://doi.org/10.1109/ICASSP39728.2021.9414219>

AUTHORS: T. -K. Hu, F. Gama, T. Chen, Z. Wang, A. Ribeiro and B. M. Sadler

HIGHLIGHT: In this work, we propose to learn decentralized controllers based solely on raw visual inputs.

982, TITLE: Recognition of Dynamic Hand Gesture Based on Mm-Wave Fmcw Radar Micro-Doppler Signatures

<https://doi.org/10.1109/ICASSP39728.2021.9414837>

AUTHORS: W. Jiang, Y. Ren, Y. Liu, Z. Wang and X. Wang

HIGHLIGHT: In this paper, we propose a convolutional neural network (CNN) for dynamic HGR based on a millimeter-wave Frequency Modulated Continuous Wave (FMCW) radar which operates at 77GHz.

983, TITLE: Dynamic Resource Optimization for Adaptive Federated Learning at the Wireless Network Edge

<https://doi.org/10.1109/ICASSP39728.2021.9414832>

AUTHORS: P. D. Lorenzo, C. Battiloro, M. Merluzzi and S. Barbarossa

HIGHLIGHT: The aim of this paper is to propose a novel dynamic resource allocation strategy for energy-efficient federated learning at the wireless network edge, with latency and learning performance guarantees.

984, TITLE: Deep Weighted MMSE Downlink Beamforming

<https://doi.org/10.1109/ICASSP39728.2021.9414561>

AUTHORS: L. Pellaco, M. Bengtsson and J. Jald'n

HIGHLIGHT: With the success of deep unfolding in trading off complexity and performance, we propose to apply deep unfolding to the WMMSE algorithm.

985, TITLE: Deep Generative Model Learning For Blind Spectrum Cartography with NMF-Based Radio Map Disaggregation

<https://doi.org/10.1109/ICASSP39728.2021.9413382>

AUTHORS: S. Shrestha, X. Fu and M. Hong

HIGHLIGHT: Our method integrates the favorable traits of model and data-driven approaches, which substantially ?shrinks? the state space.

- 986, TITLE: Mitigating Clipping Distortion in OFDM Using Deep Residual Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9413502>  
AUTHORS: M. S. Omar and X. Ma  
HIGHLIGHT: In this work, we propose a novel technique, using residual neural networks and soft clipping, to deterministically limit the peak amplitude of the signal, thus lowering its PAPR and circumventing PA distortion.
- 987, TITLE: A Low-Complexity Admm-Based Massive MIMO Detectors Via Deep Neural Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414211>  
AUTHORS: I. N. Tiba, Q. Zhang, J. Jiang and Y. Wang  
HIGHLIGHT: This paper presents a deep neural network (DNN)-based massive MIMO detection method which can overcome the above limitation.
- 988, TITLE: Real-Time Radio Modulation Classification With An LSTM Auto-Encoder  
<https://doi.org/10.1109/ICASSP39728.2021.9414351>  
AUTHORS: Z. Ke and H. Vikalo  
HIGHLIGHT: To this end, we propose a learning framework based on an LSTM denoising autoencoder designed to extract robust and stable features from the noisy received signals, and detect the underlying modulation scheme.
- 989, TITLE: Deep Active Learning Approach to Adaptive Beamforming for mmWave Initial Alignment  
<https://doi.org/10.1109/ICASSP39728.2021.9414523>  
AUTHORS: F. Sohrabi, Z. Chen and W. Yu  
HIGHLIGHT: This paper proposes a deep learning approach to the adaptive and sequential beamforming design problem for the initial access phase in a mmWave environment with a single-path channel model.
- 990, TITLE: Learning to Continuously Optimize Wireless Resource in Episodically Dynamic Environment  
<https://doi.org/10.1109/ICASSP39728.2021.9413503>  
AUTHORS: H. Sun, W. Pu, M. Zhu, X. Fu, T. -H. Chang and M. Hong  
HIGHLIGHT: We propose a continual learning (CL) framework for wireless systems, which can incrementally adapt the learning models to the new episodes, without forgetting models learned from the previous episodes.
- 991, TITLE: Adaptive Contention Window Design Using Deep Q-Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9414805>  
AUTHORS: A. Kumar, G. Verma, C. Rao, A. Swami and S. Segarra  
HIGHLIGHT: More precisely, our goal is to design an intelligent node that can dynamically adapt its minimum CW (MCW) parameter to maximize a network-level utility knowing neither the MCWs of other nodes nor how these change over time.
- 992, TITLE: On Information Asymmetry in Online Reinforcement Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9413968>  
AUTHORS: E. Tampubolon, H. Ceribasi and H. Boche  
HIGHLIGHT: In this work, we study the system of two interacting non-cooperative Q-learning agents, where one agent has the privilege of observing the other's actions.
- 993, TITLE: Jamming Strategy Generation for Hidden Communication Modes Via Graph Convolution Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414154>  
AUTHORS: F. Kong, Q. Li and H. Shao  
HIGHLIGHT: In this work, we focus on the less studied hidden mode jamming problem.
- 994, TITLE: Contrastive Self-Supervised Learning for Wireless Power Control  
<https://doi.org/10.1109/ICASSP39728.2021.9413621>  
AUTHORS: N. Naderializadeh  
HIGHLIGHT: We propose a new approach for power control in wireless networks using self-supervised learning.
- 995, TITLE: Measure-Transformed Covariance Test for Robust Spectrum Sensing  
<https://doi.org/10.1109/ICASSP39728.2021.9414592>  
AUTHORS: Y. Sorek and K. Todros  
HIGHLIGHT: In this paper, we develop a new robust spectrum sensing method for MIMO cognitive radios in the presence of heavy-tailed noise.
- 996, TITLE: Searching for Anomalies with Multiple Plays under Delay and Switching Costs

<https://doi.org/10.1109/ICASSP39728.2021.9413537>

AUTHORS: T. Lambez and K. Cohen

HIGHLIGHT: We develop a policy, dubbed consecutive controlled sensing (CCS), to achieve this goal.

997, TITLE: Robust estimation of high-order phase dynamics using Variational Bayes inference

<https://doi.org/10.1109/ICASSP39728.2021.9415028>

AUTHORS: F. Fabozzi, S. Bidon and S. Roche

HIGHLIGHT: In this paper, we propose a new nonlinear phase estimator to obtain more robust tracks.

998, TITLE: Robust PCA Through Maximum Correntropy Power Iterations

<https://doi.org/10.1109/ICASSP39728.2021.9414347>

AUTHORS: J. P. Chereau, B. Scalzo and D. P. Mandic

HIGHLIGHT: To this end, we introduce a robust formulation of PCA based on the maximum correntropy criterion (MCC).

999, TITLE: Score-Based Change Detection For Gradient-Based Learning Machines

<https://doi.org/10.1109/ICASSP39728.2021.9414085>

AUTHORS: L. Liu, J. Salmon and Z. Harchaoui

HIGHLIGHT: We present a generic score-based change detection method that can detect a change in any number of components of a machine learning model trained via empirical risk minimization.

1000, TITLE: Super-Resolution Of Periodic Signals From Short Sequences Of Samples

<https://doi.org/10.1109/ICASSP39728.2021.9413593>

AUTHORS: M. W. Rupniewski

HIGHLIGHT: In this paper, we propose a novel algorithm which does not require the signal to be bandlimited and it can cope with additive noise in the samples.

1001, TITLE: Quickest Change Detection With Time Inconsistent Anticipatory Agents In Cyber-Physical Systems

<https://doi.org/10.1109/ICASSP39728.2021.9414836>

AUTHORS: V. Krishnamurthy

HIGHLIGHT: We show that the interaction between anticipatory agents and sequential quickest detection results in unusual (nonconvex) structure of the quickest change detection policy.

1002, TITLE: Treatment Effect Estimation Using Invariant Risk Minimization

<https://doi.org/10.1109/ICASSP39728.2021.9413705>

AUTHORS: A. Shah, K. Ahuja, K. Shanmugam, D. Wei, K. R. Varshney and A. Dhurandhar

HIGHLIGHT: In this work, we propose a new way to estimate the ITE using the domain generalization framework of invariant risk minimization (IRM).

1003, TITLE: An F-Test for Polynomial Frequency Modulation

<https://doi.org/10.1109/ICASSP39728.2021.9414209>

AUTHORS: K. Blanchette, W. Burr and G. Takahara

HIGHLIGHT: We propose a semi-parametric multitaper F -test for the detection of line components which have been modulated by a polynomial of a given degree.

1004, TITLE: Quickest Joint Detection and Classification of Faults in Statistically Periodic Processes

<https://doi.org/10.1109/ICASSP39728.2021.9414101>

AUTHORS: T. Banerjee, S. Padhy, A. Taha and E. John

HIGHLIGHT: An algorithm is proposed to detect and classify a change in the distribution of a stochastic process that has periodic statistical behavior.

1005, TITLE: An Asymptotically Pointwise Optimal Procedure For Sequential Joint Detection And Estimation

<https://doi.org/10.1109/ICASSP39728.2021.9414365>

AUTHORS: D. Reinhard, M. Fau? and A. M. Zoubir

HIGHLIGHT: To solve the optimal stopping problem, we propose an asymptotically pointwise optimal (APO) stopping rule, i.e., a stopping rule that is optimal when the tolerated detection and estimation errors tend to zero.

1006, TITLE: Locally Optimal Detection of Stochastic Targeted Universal Adversarial Perturbations

<https://doi.org/10.1109/ICASSP39728.2021.9413866>

AUTHORS: A. Goel and P. Moulin

**HIGHLIGHT:** In this paper, we derive the locally optimal generalized likelihood ratio test based detector for detecting stochastic targeted universal adversarial perturbations to a classifier's input.

1007, **TITLE:** A Decentralized Variance-Reduced Method for Stochastic Optimization Over Directed Graphs

<https://doi.org/10.1109/ICASSP39728.2021.9413600>

**AUTHORS:** M. I. Qureshi, R. Xin, S. Kar and U. A. Khan

**HIGHLIGHT:** In this paper, we propose a decentralized first-order stochastic optimization method Push-SAGA for finite-sum minimization over a strongly connected directed graph.

1008, **TITLE:** On Distributed Composite Tests with Dependent Observations in WSN

<https://doi.org/10.1109/ICASSP39728.2021.9415115>

**AUTHORS:** J. A. Maya and L. Rey Vega

**HIGHLIGHT:** As the sensors observe statistically spatial dependent samples, which makes difficult the implementation of fully distributed detection procedures, we propose a simpler algorithm for making a decision about the true hypothesis.

1009, **TITLE:** Byzantine-Resilient Decentralized TD Learning with Linear Function Approximation

<https://doi.org/10.1109/ICASSP39728.2021.9413992>

**AUTHORS:** Z. Wu, H. Shen, T. Chen and Q. Ling

**HIGHLIGHT:** This paper considers the policy evaluation problem in reinforcement learning with agents of a decentralized and directed network.

1010, **TITLE:** On The Effect of Spatial Correlation on Distributed Energy Detection of a Stochastic Process

<https://doi.org/10.1109/ICASSP39728.2021.9415097>

**AUTHORS:** J. A. Maya and L. Rey Vega

**HIGHLIGHT:** As those parameters can be easily estimated at each sensor node, we propose a modified generalized likelihood ratio test (GLRT).

1011, **TITLE:** Provably Fast Asynchronous And Distributed Algorithms For Pagerank Centrality Computation

<https://doi.org/10.1109/ICASSP39728.2021.9413643>

**AUTHORS:** Y. He and H. -T. Wai

**HIGHLIGHT:** We study asynchronous and distributed algorithms which are operated by aggregating information from local neighbors iteratively.

1012, **TITLE:** Decentralized Optimization Over Noisy, Rate-Constrained Networks: How We Agree By Talking About How We Disagree

<https://doi.org/10.1109/ICASSP39728.2021.9413527>

**AUTHORS:** R. Saha, S. Rini, M. Rao and A. Goldsmith

**HIGHLIGHT:** We develop a novel algorithm for this scenario: Decentralized Lazy Mirror Descent with Differential Exchanges (DLMD-DiffEx), which guarantees convergence of the local estimates to the optimal solution.

1013, **TITLE:** A Multiple Access Channel Game Using Latency Metric

<https://doi.org/10.1109/ICASSP39728.2021.9413529>

**AUTHORS:** A. Garnaev, A. Petropulu and W. Trappe

**HIGHLIGHT:** The paper considers a multi-access channel scenario, where several users communicate with a base station, and investigates power allocation is a game-theoretic framework.

1014, **TITLE:** Linear Computation Coding

<https://doi.org/10.1109/ICASSP39728.2021.9414317>

**AUTHORS:** R. R. Müller, B. Gode and A. Breyhi

**HIGHLIGHT:** For linear functions, we present an algorithm to reduce the computational cost of multiplying an arbitrary given matrix with an unknown vector.

1015, **TITLE:** Spectral Folding And Two-Channel Filter-Banks On Arbitrary Graphs

<https://doi.org/10.1109/ICASSP39728.2021.9414066>

**AUTHORS:** E. Pavez, B. Girault, A. Ortega and P. A. Chou

**HIGHLIGHT:** In this paper we extend this theory to arbitrary graphs and positive semi-definite variation operators.

1016, **TITLE:** Sparse Time-Frequency Representation Via Atomic Norm Minimization

<https://doi.org/10.1109/ICASSP39728.2021.9414921>

**AUTHORS:** T. Kusano, K. Yatabe and Y. Oikawa

**HIGHLIGHT:** In this paper, we propose a method of estimating a sparse T-F representation using atomic norm.

1017, TITLE: Message Transmission Over Rapidly Time-Varying Channels

<https://doi.org/10.1109/ICASSP39728.2021.9414699>

AUTHORS: A. Kaplan and V. Pohl

**HIGHLIGHT:** This paper presents a new method on transmitting data over linear time-variant (LTV) channels when no channel state estimation (CSI) is accessible.

1018, TITLE: A Robust and Efficient Multi-Scale Seasonal-Trend Decomposition

<https://doi.org/10.1109/ICASSP39728.2021.9413939>

AUTHORS: L. Yang, Q. Wen, B. Yang and L. Sun

**HIGHLIGHT:** In this paper, we propose a general and efficient multi-scale seasonal-trend decomposition algorithm for time series with multiple seasonality.

1019, TITLE: Noise-Assisted Multivariate Variational Mode Decomposition

<https://doi.org/10.1109/ICASSP39728.2021.9413367>

AUTHORS: C. A. Zisou, G. K. Apostolidis and L. J. Hadjileontiadis

**HIGHLIGHT:** This work combines MVMD with the noise injection paradigm to propose an efficient alternative for both VMD and MVMD, i.e., the noise-assisted MVMD (NA-MVMD), that aims at relaxing the requirement of presetting K, as well as improving the quality of the resulting decomposition.

1020, TITLE: Approximate Weighted C R Coded Matrix Multiplication

<https://doi.org/10.1109/ICASSP39728.2021.9413800>

AUTHORS: N. Charalambides, M. Pilanci and A. O. Hero

**HIGHLIGHT:** To enable distributed low rank approximation, we generalize the approximate matrix CR-multiplication to accommodate weighted block sampling, and we introduce a weighted coded matrix multiplication method.

1021, TITLE: Periodic Signal Denoising: An Analysis-Synthesis Framework Based on Ramanujan Filter Banks and Dictionaries

<https://doi.org/10.1109/ICASSP39728.2021.9413689>

AUTHORS: P. Kulkarni and P. P. Vaidyanathan

**HIGHLIGHT:** This paper proposes to use a hybrid analysis-synthesis framework for denoising discrete-time periodic signals.

1022, TITLE: Compressive Signal Recovery Under Sensing Matrix Errors Combined With Unknown Measurement Gains

<https://doi.org/10.1109/ICASSP39728.2021.9413470>

AUTHORS: J. Vora and A. Rajwade

**HIGHLIGHT:** We propose an alternating minimisation algorithm for on-the-fly signal recovery in the case when errors (a) and (b) occur jointly.

1023, TITLE: Grid Optimization for Matrix-Based Source Localization Under Inhomogeneous Sensor Topology

<https://doi.org/10.1109/ICASSP39728.2021.9414100>

AUTHORS: H. Sun and J. Chen

**HIGHLIGHT:** This paper finds that the localization performance degrades when the spatial pattern of the sensors is highly non-uniform, and the uniform grid formation is only a suboptimal solution.

1024, TITLE: MSR-GAN: Multi-Segment Reconstruction via Adversarial Learning

<https://doi.org/10.1109/ICASSP39728.2021.9414895>

AUTHORS: M. Zehni and Z. Zhao

**HIGHLIGHT:** Thus, we propose to approximate it using Gumbel-Softmax reparametrization trick.

1025, TITLE: Count Sketch with Zero Checking: Efficient Recovery of Heavy Components

<https://doi.org/10.1109/ICASSP39728.2021.9413853>

AUTHORS: G. Zhou and Z. Tian

**HIGHLIGHT:** In this paper, we carefully analyze count sketch and illustrate that its default recovery method, namely median filtering, has a distinct error pattern of reporting false positives.

1026, TITLE: Numerical Solution of Stochastic Differential Equations in Stiefel Manifolds via Tangent Space Parametrization

<https://doi.org/10.1109/ICASSP39728.2021.9413353>

AUTHORS: Z. Wang and V. Solo

**HIGHLIGHT:** Here we achieve this by extending the so-called 'tangent space parameterization' (TaSP) for ODEs to SDEs.

- 1027, TITLE: On The Accuracy Limit of Joint Time-Delay/Doppler/Acceleration Estimation with a Band-Limited Signal  
<https://doi.org/10.1109/ICASSP39728.2021.9414270>  
 AUTHORS: H. McPhee, L. Ortega, J. Vil?-Valls and E. Chaumette  
 HIGHLIGHT: Considering a generic band-limited signal formulation, in this contribution we derive a new closed-form Cram?-r-Rao bound (CRB) expression for joint time-delay/Doppler/acceleration estimation.
- 1028, TITLE: Automatic Order Selection in Autoregressive Modeling with Application in EEG Sleep-Stage Classification  
<https://doi.org/10.1109/ICASSP39728.2021.9414795>  
 AUTHORS: F. Nassif and S. Beheshti  
 HIGHLIGHT: This paper investigates the order selection problem for autoregressive models from a new perspective.
- 1029, TITLE: New Variants of DFA Based on Loess and Lowess Methods: Generalization of the Detrending Moving Average  
<https://doi.org/10.1109/ICASSP39728.2021.9414216>  
 AUTHORS: B. Berthelot, E. Grivel and P. Legrand  
 HIGHLIGHT: In this paper, we recall the main principles of some of these methods, provide explanations on the behaviours of the algorithms and analyze the relevance of new variants based on the Savitzky-Golay filter, also known as the LOESS approach, and the LOWESS.
- 1030, TITLE: Parameter Estimation for Student?s t VAR Model with Missing Data  
<https://doi.org/10.1109/ICASSP39728.2021.9414223>  
 AUTHORS: R. Zhou, J. Liu, S. Kumar and D. P. Palomar  
 HIGHLIGHT: In this paper, we propose an algorithmic framework to estimate the parameters of a VAR model with heavy-tailed Student?s t distributed innovations from incomplete data based on the stochastic approximation expectation maximization (SAEM) algorithm coupled with a Markov Chain Monte Carlo (MCMC) procedure.
- 1031, TITLE: Fast and Robust ADMM for Blind Super-Resolution  
<https://doi.org/10.1109/ICASSP39728.2021.9415003>  
 AUTHORS: Y. Ran and W. Dai  
 HIGHLIGHT: To do so, we adapt an operator splitting approach ADMM and combine it with a novel preconditioning scheme.
- 1032, TITLE: Nonstationary Portfolios: Diversification in the Spectral Domain  
<https://doi.org/10.1109/ICASSP39728.2021.9413769>  
 AUTHORS: B. Scalzo, A. Arroyo, L. Stankovic and D. P. Mandic  
 HIGHLIGHT: To this end, we reformulate the portfolio optimization problem in the spectral domain to cater for the nonstationarity inherent to asset price movements and, in this way, allow for optimal capital allocations to be time-varying.
- 1033, TITLE: A Tyler-Type Estimator of Location and Scatter Leveraging Riemannian Optimization  
<https://doi.org/10.1109/ICASSP39728.2021.9414974>  
 AUTHORS: A. Collas, F. Bouchard, A. Breloy, C. Ren, G. Ginolhac and J. . -P. Ovarlez  
 HIGHLIGHT: In this paper, we propose a stable algorithm based on Riemannian optimization for this problem.
- 1034, TITLE: Statistical Properties of a Modified Welch Method That Uses Sample Percentiles  
<https://doi.org/10.1109/ICASSP39728.2021.9415074>  
 AUTHORS: F. Schwock and S. Abadi  
 HIGHLIGHT: We present and analyze an alternative, more robust approach to the Welch?s overlapped segment averaging (WOSA) spectral estimator.
- 1035, TITLE: Switched Hawkes Processes  
<https://doi.org/10.1109/ICASSP39728.2021.9413812>  
 AUTHORS: N. Nadagouda and M. A. Davenport  
 HIGHLIGHT: In this paper, we consider a switched Hawkes process which can be used to model systems in which the parameters of the process dynamically change depending on some (known) external state.
- 1036, TITLE: An Adaptive Regularization Approach to Portfolio Optimization  
<https://doi.org/10.1109/ICASSP39728.2021.9413865>  
 AUTHORS: T. Ballal, A. S. Abdelrahman, A. H. Muqaibel and T. Y. Al-Naffouri  
 HIGHLIGHT: In this work, we address the regularization issue from a different perspective.

- 1037, TITLE: Active Estimation From Multimodal Data  
<https://doi.org/10.1109/ICASSP39728.2021.9414772>  
AUTHORS: A. Mukherjee, A. Tager, P. -Y. Chen and P. Das  
HIGHLIGHT: The paper considers the problem of estimating a covariate parameter shared by multiple statistical models.
- 1038, TITLE: Network Classifiers Based on Social Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9414126>  
AUTHORS: V. Bordinon, S. Vlaski, V. Matta and A. H. Sayed  
HIGHLIGHT: This work proposes a new way of combining independently trained classifiers over space and time.
- 1039, TITLE: Bayes-Optimal Methods for Finding the Source of a Cascade  
<https://doi.org/10.1109/ICASSP39728.2021.9414392>  
AUTHORS: A. Sridhar and H. V. Poor  
HIGHLIGHT: We study the problem of estimating the source of a network cascade.
- 1040, TITLE: Private Wireless Federated Learning with Anonymous Over-the-Air Computation  
<https://doi.org/10.1109/ICASSP39728.2021.9413624>  
AUTHORS: B. Hasircioglu and D. G?nd?z  
HIGHLIGHT: In the wireless FL scenario, we show that the privacy of the system can be boosted by exploiting over-the-air computation (OAC) and anonymizing the transmitting devices.
- 1041, TITLE: Scalable Multilevel Quantization for Distributed Detection  
<https://doi.org/10.1109/ICASSP39728.2021.9414032>  
AUTHORS: G. G?l and M. Ba?ler  
HIGHLIGHT: A scalable algorithm is derived for multilevel quantization of sensor observations in distributed sensor networks, which consist of a number of sensors transmitting a summary information of their observations to the fusion center for a final decision.
- 1042, TITLE: Stability of Algebraic Neural Networks to Small Perturbations  
<https://doi.org/10.1109/ICASSP39728.2021.9414604>  
AUTHORS: A. Parada-Mayorga and A. Ribeiro  
HIGHLIGHT: In this paper we study stability of AlgNNs on the framework of algebraic signal processing.
- 1043, TITLE: Resolution Limits of 20 Questions Search Strategies for Moving Targets  
<https://doi.org/10.1109/ICASSP39728.2021.9414680>  
AUTHORS: L. Zhou and A. Hero  
HIGHLIGHT: We are interested in the minimal resolution of any non-adaptive searching procedure with a finite number of queries and derive approximations to this optimal resolution via the second-order asymptotic analysis.
- 1044, TITLE: Gramian-Based Adaptive Combination Policies for Diffusion Learning Over Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414449>  
AUTHORS: Y. E. Erginbas, S. Vlaski and A. H. Sayed  
HIGHLIGHT: This paper presents an adaptive combination strategy for distributed learning over diffusion networks.
- 1045, TITLE: Graph-Adaptive Incremental Learning Using an Ensemble of Gaussian Process Experts  
<https://doi.org/10.1109/ICASSP39728.2021.9413970>  
AUTHORS: K. D. Polyzos, Q. Lu and G. B. Giannakis  
HIGHLIGHT: To allow for uncertainty quantification, which is of utmost importance in safety-critical applications, this work tackles the SSL task in a Gaussian process (GP) based Bayesian framework to propagate the distribution of nonparametric function estimates.
- 1046, TITLE: Fast Decentralized Linear Functions Via Successive Graph Shift Operators  
<https://doi.org/10.1109/ICASSP39728.2021.9415098>  
AUTHORS: S. Mollaebrahim, D. Romero and B. Beferull-Lozano  
HIGHLIGHT: In contrast, this paper develops a decentralized method to compute linear transformations in a small number of iterations.
- 1047, TITLE: Online Learning of Time-Varying Signals and Graphs  
<https://doi.org/10.1109/ICASSP39728.2021.9415029>  
AUTHORS: S. Sardellitti, S. Barbarossa and P. Di Lorenzo



**HIGHLIGHT:** The aim of this paper is to propose a method for online learning of time-varying graphs from noisy observations of smooth graph signals collected over the vertices.

1048, **TITLE:** Kernel Regression on Graphs in Random Fourier Features Space

<https://doi.org/10.1109/ICASSP39728.2021.9414951>

**AUTHORS:** V. R. M. Elias, V. C. Gogineni, W. A. Martins and S. Werner

**HIGHLIGHT:** This work proposes an efficient batch-based implementation for kernel regression on graphs (KRG) using random Fourier features (RFF) and a low-complexity online implementation.

1049, **TITLE:** Graph-Homomorphic Perturbations for Private Decentralized Learning

<https://doi.org/10.1109/ICASSP39728.2021.9413811>

**AUTHORS:** S. Vlaski and A. H. Sayed

**HIGHLIGHT:** We propose an alternative scheme, which constructs perturbations according to a particular nullspace condition, allowing them to be invisible (to first order in the step-size) to the network centroid, while preserving privacy guarantees.

1050, **TITLE:** Variance-Constrained Learning for Stochastic Graph Neural Networks

<https://doi.org/10.1109/ICASSP39728.2021.9413751>

**AUTHORS:** Z. Gao, E. Isufi and A. Ribeiro

**HIGHLIGHT:** We propose an alternating primal-dual learning algorithm that updates the primal variable (SGNN parameters) with gradient descent and the dual variable with gradient ascent.

1051, **TITLE:** Graph Neural Network for Large-Scale Network Localization

<https://doi.org/10.1109/ICASSP39728.2021.9414520>

**AUTHORS:** W. Yan, D. Jin, Z. Lin and F. Yin

**HIGHLIGHT:** In this work, we adopt GNN for a classic but challenging nonlinear regression problem, namely the network localization.

1052, **TITLE:** Graphon and Graph Neural Network Stability

<https://doi.org/10.1109/ICASSP39728.2021.9414838>

**AUTHORS:** L. Ruiz, Z. Wang and A. Ribeiro

**HIGHLIGHT:** We analyze GNN stability using kernel objects called graphons.

1053, **TITLE:** Graph Neural Networks for Decentralized Controllers

<https://doi.org/10.1109/ICASSP39728.2021.9414563>

**AUTHORS:** F. Gama, E. Tolstaya and A. Ribeiro

**HIGHLIGHT:** In this paper, we propose a framework using graph neural networks (GNNs) to learn decentralized controllers from data.

1054, **TITLE:** Nonlinear State-Space Generalizations of Graph Convolutional Neural Networks

<https://doi.org/10.1109/ICASSP39728.2021.9414672>

**AUTHORS:** L. Ruiz, F. Gama, A. Ribeiro and E. Isufi

**HIGHLIGHT:** In this work, we approach GCNNs from a state-space perspective revealing that the graph convolutional module is a minimalistic linear state-space model, in which the state update matrix is the graph shift operator.

1055, **TITLE:** Wide and Deep Graph Neural Networks with Distributed Online Learning

<https://doi.org/10.1109/ICASSP39728.2021.9415046>

**AUTHORS:** Z. Gao, A. Ribeiro and F. Gama

**HIGHLIGHT:** This paper proposes the Wide and Deep GNN (WD-GNN), a novel architecture that can be easily updated with distributed online learning mechanisms.

1056, **TITLE:** Design of Graph Signal Sampling Matrices for Arbitrary Signal Subspaces

<https://doi.org/10.1109/ICASSP39728.2021.9414240>

**AUTHORS:** J. Hara, K. Yamada, S. Ono and Y. Tanaka

**HIGHLIGHT:** We propose a design method of sampling matrices for graph signals that guarantees perfect recovery for arbitrary graph signal subspaces.

1057, **TITLE:** Graph Signal Denoising Using Nested-Structured Deep Algorithm Unrolling

<https://doi.org/10.1109/ICASSP39728.2021.9414093>

**AUTHORS:** M. Nagahama, K. Yamada, Y. Tanaka, S. H. Chan and Y. C. Eldar

**HIGHLIGHT:** In this paper, we propose a deep algorithm unrolling (DAU) based on a variant of the alternating direction method of multiplier (ADMM) called Plug-and-Play ADMM (PnP-ADMM) for denoising of signals on graphs.

1058, **TITLE:** Identifying First-Order Lowpass Graph Signals Using Perron Frobenius Theorem

<https://doi.org/10.1109/ICASSP39728.2021.9415031>

**AUTHORS:** Y. He and H. -T. Wai

**HIGHLIGHT:** Our aim is to determine if the graph filter generating the graph signals is first-order lowpass without knowing the graph topology.

1059, **TITLE:** Graph Signal Denoising Via Unrolling Networks

<https://doi.org/10.1109/ICASSP39728.2021.9415073>

**AUTHORS:** S. Chen and Y. C. Eldar

**HIGHLIGHT:** We propose an interpretable graph neural network framework to denoise single or multiple noisy graph signals.

1060, **TITLE:** Adaptive Subsampling of Multidomain Signals with Product Graphs

<https://doi.org/10.1109/ICASSP39728.2021.9413459>

**AUTHORS:** T. Gnassounou, P. Humbert and L. Oudre

**HIGHLIGHT:** In this paper, we propose an adaptive subsampling method for multidomain signals based on the constrained learning of a product graph.

1061, **TITLE:** Robust Graph-Filter Identification with Graph Denoising Regularization

<https://doi.org/10.1109/ICASSP39728.2021.9414909>

**AUTHORS:** S. Rey and A. G. Marques

**HIGHLIGHT:** Tailored to those setups, this paper presents a robust formulation for the problem of graph-filter identification from input-output observations.

1062, **TITLE:** Fast and Provable Robust PCA VIA Normalized Coherence Pursuit

<https://doi.org/10.1109/ICASSP39728.2021.9414012>

**AUTHORS:** M. Rahmani and P. Li

**HIGHLIGHT:** We study the Innovation Values computed by the Innovation Search algorithm under a quadratic cost function and it is proved that Innovation Values with the new cost function are equivalent to Leverage Scores.

1063, **TITLE:** Aligning Sets of Temporal Signals with Riemannian Geometry and Koopman Operator

<https://doi.org/10.1109/ICASSP39728.2021.9413729>

**AUTHORS:** O. Rahamim and R. Talmon

**HIGHLIGHT:** In this paper, we consider the problem of aligning data sets of short temporal signals without any a-priori known correspondence.

1064, **TITLE:** Weight Identification Through Global Optimization in a New Hysteretic Neural Network Model

<https://doi.org/10.1109/ICASSP39728.2021.9413383>

**AUTHORS:** E. Leroy, A. Marmin, M. Castella and L. Duval

**HIGHLIGHT:** In this contribution, we propose to consider new nonlinear activation functions, whose outputs depend both from the current and past inputs through a hysteresis effect.

1065, **TITLE:** Multiview Variational Graph Autoencoders for Canonical Correlation Analysis

<https://doi.org/10.1109/ICASSP39728.2021.9414466>

**AUTHORS:** Y. Kaloga, P. Borgnat, S. P. Chepuri, P. Abry and A. Habrard

**HIGHLIGHT:** We present a novel multiview canonical correlation analysis model based on a variational approach.

1066, **TITLE:** Cognitive Memory Constrained Human Decision Making based on Multi-source Information

<https://doi.org/10.1109/ICASSP39728.2021.9413745>

**AUTHORS:** B. Geng, Q. Chen and P. K. Varshney

**HIGHLIGHT:** The objective of this work is to study how humans make decisions based on internal and external sources of information under cognitive memory limitations.

1067, **TITLE:** Binary Control and Digital-to-Analog Conversion Using Composite NUV Priors and Iterative Gaussian Message Passing

<https://doi.org/10.1109/ICASSP39728.2021.9413677>

**AUTHORS:** R. Keusch, H. Malmberg and H. -A. Loeliger

**HIGHLIGHT:** The paper proposes a new method to determine a binary control signal for an analog linear system such that the state, or some output, of the system follows a given target trajectory.

1068, **TITLE:** Outlier-Robust Kernel Hierarchical-Optimization RLS on a Budget with Affine Constraints

<https://doi.org/10.1109/ICASSP39728.2021.9413415>

**AUTHORS:** K. Slavakis and M. Yukawa

**HIGHLIGHT:** This paper introduces a non-parametric learning framework to combat outliers in online, multi-output, and nonlinear regression tasks.

1069, **TITLE:** Adaptive Real-Time Filter for Partially-Observed Boolean Dynamical Systems

<https://doi.org/10.1109/ICASSP39728.2021.9413485>

**AUTHORS:** M. Imani and S. F. Ghoreishi

**HIGHLIGHT:** In this manuscript, we present a real-time joint state and parameter estimation framework for POBDS.

1070, **TITLE:** Improving the Energy-Efficiency of a Kalman Filter Using Unreliable Memories

<https://doi.org/10.1109/ICASSP39728.2021.9413430>

**AUTHORS:** J. Kern, E. Dupraz, A. A?ssa-El-Bey and F. Leduc-Primeau

**HIGHLIGHT:** The degradation in estimation performance caused by the noise in the memory is theoretically investigated. Updated equations are then developed for the Kalman filter, taking into account the new source of noise from the unreliable memory.

1071, **TITLE:** Parallel Iterated Extended and Sigma-Point Kalman Smoothers

<https://doi.org/10.1109/ICASSP39728.2021.9413364>

**AUTHORS:** F. Yaghoobi, A. Corenflos, S. Hassan and S. S?rkk?

**HIGHLIGHT:** In this paper, we present a set of parallel formulas that replace the existing sequential ones in order to achieve lower time (span) complexity.

1072, **TITLE:** Wiener Filter on Meet/Join Lattices

<https://doi.org/10.1109/ICASSP39728.2021.9415055>

**AUTHORS:** B. Seifert, C. Wendler and M. P?schel

**HIGHLIGHT:** In this paper we extend DLSP with Wiener filtering for denoising and demonstrate it on two prototypical applications.

1073, **TITLE:** Learning Bollob?s-Riordan Graphs Under Partial Observability

<https://doi.org/10.1109/ICASSP39728.2021.9414217>

**AUTHORS:** M. Cirillo, V. Matta and A. H. Sayed

**HIGHLIGHT:** In this work, we discover that, over first-order vector autoregressive systems with a stable Laplacian combination matrix, graph learning is achievable under partial observability, when the network topology is drawn according to a popular preferential attachment model known as the Bollob?s-Riordan model.

1074, **TITLE:** Learning Sparse Graph Laplacian with K Eigenvector Prior via Iterative Glasso and Projection

<https://doi.org/10.1109/ICASSP39728.2021.9414693>

**AUTHORS:** S. Bagheri, G. Cheung, A. Ortega and F. Wang

**HIGHLIGHT:** In this paper, given an empirical covariance matrix  $\mathbf{\bar{C}}$  computed from data as input, we consider an eigen-structural assumption on the graph Laplacian matrix  $L$ : the first  $K$  eigenvectors of  $L$  are pre-selected, e.g., based on domain-specific criteria, and the remaining eigenvectors are then learned from data.

1075, **TITLE:** Learning Mixed Membership from Adjacency Graph Via Systematic Edge Query: Identifiability and Algorithm

<https://doi.org/10.1109/ICASSP39728.2021.9413541>

**AUTHORS:** S. Ibrahim and X. Fu

**HIGHLIGHT:** A lightweight scalable algorithm is proposed, and its performance characterizations are presented.

1076, **TITLE:** Convergence Analysis of the Graph-Topology-Inference Kernel LMS Algorithm

<https://doi.org/10.1109/ICASSP39728.2021.9413693>

**AUTHORS:** M. Moscu, R. Borsoi and C. Richard

**HIGHLIGHT:** In this paper, we present a convergence analysis of the graph-topology-inference KLMS algorithm.

1077, **TITLE:** An Efficient Alternating Direction Method for Graph Learning from Smooth Signals

<https://doi.org/10.1109/ICASSP39728.2021.9414791>

**AUTHORS:** X. Wang, C. Yao, H. Lei and A. M. -C. So

**HIGHLIGHT:** In this paper, we cast the graph learning formulation as a nonsmooth, strictly convex optimization problem and develop an efficient alternating direction method of multipliers to solve it.

1078, **TITLE:** Topological Volterra Filters  
<https://doi.org/10.1109/ICASSP39728.2021.9414275>  
**AUTHORS:** G. Leus, M. Yang, M. Coutino and E. Isufi  
**HIGHLIGHT:** We apply TVFs to inverse filtering and recommender systems.

1079, **TITLE:** Network Topology Inference with Graphon Spectral Penalties  
<https://doi.org/10.1109/ICASSP39728.2021.9414266>  
**AUTHORS:** T. M. Roddenberry, M. Navarro and S. Segarra  
**HIGHLIGHT:** We present the cases where the graphon model is assumed to be known and the more practical setting where the relevant features of the model are inferred from auxiliary network observations.

1080, **TITLE:** Network Topology Change-Point Detection from Graph Signals with Prior Spectral Signatures  
<https://doi.org/10.1109/ICASSP39728.2021.9413857>  
**AUTHORS:** C. Kaushik, T. M. Roddenberry and S. Segarra  
**HIGHLIGHT:** We consider the problem of sequential graph topology change-point detection from graph signals.

1081, **TITLE:** Online Time-Varying Topology Identification Via Prediction-Correction Algorithms  
<https://doi.org/10.1109/ICASSP39728.2021.9415053>  
**AUTHORS:** A. Natali, M. Coutino, E. Isufi and G. Leus  
**HIGHLIGHT:** In this paper, we address the problem of dynamic topology identification by building on recent results from time-varying optimization, devising a general-purpose online algorithm operating in non-stationary environments.

1082, **TITLE:** Graph Learning Under Spectral Sparsity Constraints  
<https://doi.org/10.1109/ICASSP39728.2021.9413561>  
**AUTHORS:** B. Subbareddy, A. Siripuram and J. Zhang  
**HIGHLIGHT:** We propose a new signal processing based inference model and a new learning criterion that allow for wideband frequency variation in the data and derive an algorithm for graph inference.

1083, **TITLE:** A Graph Learning Algorithm Based On Gaussian Markov Random Fields And Minimax Concave Penalty  
<https://doi.org/10.1109/ICASSP39728.2021.9413850>  
**AUTHORS:** T. Koyakumar, M. Yukawa, E. Pavez and A. Ortega  
**HIGHLIGHT:** This paper presents a graph learning framework to produce sparse and accurate graphs from network data.

1084, **TITLE:** Figlearn: Filter and Graph Learning Using Optimal Transport  
<https://doi.org/10.1109/ICASSP39728.2021.9413778>  
**AUTHORS:** M. Minder, Z. Farsijani, D. Shah, M. E. Gheche and P. Frossard  
**HIGHLIGHT:** We hence introduce a novel graph signal processing framework for jointly learning the graph and its generating filter from signal observations.

1085, **TITLE:** Construction of Unit-Norm Tight Frame Based Preconditioner for Sparse Coding  
<https://doi.org/10.1109/ICASSP39728.2021.9414404>  
**AUTHORS:** H. Bai, C. Hong and X. Li  
**HIGHLIGHT:** In this paper, we investigate the problem of constructing suitable preconditioner to improve the performance of the SR system.

1086, **TITLE:** Sparse High-Order Portfolios Via Proximal Dca And Sca  
<https://doi.org/10.1109/ICASSP39728.2021.9414048>  
**AUTHORS:** J. Wang, Z. Deng, T. Zheng and A. M. -C. So  
**HIGHLIGHT:** In this paper, we study the cardinality constrained mean-variance-skewness-kurtosis (MVSKC) model for sparse high-order portfolio optimization.

1087, **TITLE:** A Convex Penalty for Block-Sparse Signals with Unknown Structures  
<https://doi.org/10.1109/ICASSP39728.2021.9414340>  
**AUTHORS:** H. Kuroda, D. Kitahara and A. Hirabayashi  
**HIGHLIGHT:** We propose a novel convex penalty for block-sparse signals whose block partitions are unknown a priori.

1088, TITLE: Event-Driven Modulo Sampling  
<https://doi.org/10.1109/ICASSP39728.2021.9414152>  
AUTHORS: D. Florescu, F. Krahmer and A. Bhandari  
HIGHLIGHT: We propose a cascade model comprising a modulo non-linearity in series with an integrate-and-fire (IF) event-driven encoder.

1089, TITLE: No Relaxation: Guaranteed Recovery of Finite-Valued Signals from Undersampled Measurements  
<https://doi.org/10.1109/ICASSP39728.2021.9413945>  
AUTHORS: P. Sarangi and P. Pal  
HIGHLIGHT: We develop a new computationally efficient decoding algorithm that can operate at the optimal downsampling factor under mild conditions on the filter.

1090, TITLE: Error Estimates in Second-Order Continuous-Time Sigma-Delta Modulators  
<https://doi.org/10.1109/ICASSP39728.2021.9413672>  
AUTHORS: D. Surroop, P. Combes and P. Martin  
HIGHLIGHT: This paper presents a uniqueness result which states that a band-limited signal can be recovered from at least  $3B$  measurements where  $B$  is the bandwidth from the radar ambiguity function (AF).

1091, TITLE: Banraw: Band-Limited Radar Waveform Design Via Phase Retrieval  
<https://doi.org/10.1109/ICASSP39728.2021.9414781>  
AUTHORS: S. Pinilla, K. V. Mishr, B. M. Sadler and H. Arguello  
HIGHLIGHT: This paper presents a uniqueness result which states that a band-limited signal can be recovered from at least  $3B$  measurements where  $B$  is the bandwidth from the radar ambiguity function (AF).

1092, TITLE: Sub-NYQUIST Multichannel Blind Deconvolution  
<https://doi.org/10.1109/ICASSP39728.2021.9413856>  
AUTHORS: S. Mulleti, K. Lee and Y. C. Eldar  
HIGHLIGHT: We present necessary and sufficient conditions for the unique identification.

1093, TITLE: Modified Arcsine Law for One-Bit Sampled Stationary Signals with Time-Varying Thresholds  
<https://doi.org/10.1109/ICASSP39728.2021.9414992>  
AUTHORS: A. Eamazi, F. Yeganegi and M. Soltanalian  
HIGHLIGHT: For the first time in the literature, this paper introduces an approach to extending the arcsine law to the case where one-bit ADCs apply time-varying thresholds.

1094, TITLE: Near-Optimal Resampling in Particle Filters Using the Ising Energy Model  
<https://doi.org/10.1109/ICASSP39728.2021.9413633>  
AUTHORS: M. T. Rahman, M. Javad-Kalbasi and S. Valaee  
HIGHLIGHT: Instead of utilizing resampling procedures that rely on asymptotic convergence properties, we show that intelligently selecting and replicating a set of samples can better represent the posterior approximation and improve the overall performance of the PF.

1095, TITLE: Time-Domain Concentration and Approximation of Computable Bandlimited Signals  
<https://doi.org/10.1109/ICASSP39728.2021.9413984>  
AUTHORS: H. Boche and U. J. M. N. J.  
HIGHLIGHT: In this paper we provide a different definition that uses the time-domain concentration of the signals.

1096, TITLE: Guaranteed Reconstruction from Integrate-and-Fire Neurons with Alpha Synaptic Activation  
<https://doi.org/10.1109/ICASSP39728.2021.9414759>  
AUTHORS: M. Hilton, R. Alexandru and P. L. Dragotti  
HIGHLIGHT: In contrast to state-of-the-art methods we propose an IF-TEM where we employ a biologically inspired and smooth sampling kernel, the alpha synaptic function, and show that perfect reconstruction can be achieved using this kernel.

1097, TITLE: Social Learning Under Inferential Attacks  
<https://doi.org/10.1109/ICASSP39728.2021.9413642>  
AUTHORS: K. Ntemos, V. Bordinon, S. Vlaski and A. H. Sayed  
HIGHLIGHT: In this work, we consider the scenario where a subset of agents aims at driving the network beliefs to the wrong hypothesis.

1098, TITLE: Segregation in Social Networks: MARKOV Bridge Models and Estimation

<https://doi.org/10.1109/ICASSP39728.2021.9413690>

AUTHORS: V. Krishnamurthy, R. Luo and B. Nettasinghe

HIGHLIGHT: Specifically, we present a novel community-based graph model that represent segregation as a Markov bridge process.

1099, TITLE: Controlled Testing and Isolation for Suppressing Covid-19

<https://doi.org/10.1109/ICASSP39728.2021.9414482>

AUTHORS: K. Cohen and A. Leshem

HIGHLIGHT: We argue that controlling the spread of the virus can be done by using active feedback to control testing for infection by actively testing individuals with a high probability of being infected.

1100, TITLE: Two-Stage Graph-Constrained Group Testing: Theory and Application

<https://doi.org/10.1109/ICASSP39728.2021.9414802>

AUTHORS: S. Sihag, A. Tajer and U. Mitra

HIGHLIGHT: This paper, in contrast, formalizes an adaptive, two-stage framework that requires  $T(kM^2 \log(p/k))$  tests, that is, a factor  $k$  smaller than that of the one-stage (non-adaptive) frameworks.

1101, TITLE: Unveiling Anomalous Nodes Via Random Sampling and Consensus on Graphs

<https://doi.org/10.1109/ICASSP39728.2021.9414953>

AUTHORS: V. N. Ioannidis, D. Berberidis and G. B. Giannakis

HIGHLIGHT: The present paper develops a graph-based sampling and consensus (GraphSAC) approach to effectively detect anomalous nodes in large-scale graphs.

1102, TITLE: Estimating Fiedler Value on Large Networks Based on Random Walk Observations

<https://doi.org/10.1109/ICASSP39728.2021.9413713>

AUTHORS: A. Reiffers-Masson, T. Chonavel and Y. Hayel

HIGHLIGHT: In this paper, we describe an iterative scheme which is able to estimate the Fiedler value of a network when the topology is initially unknown.

1103, TITLE: Orthogonality and Zero DC Tradeoffs in Biorthogonal Graph Filterbanks

<https://doi.org/10.1109/ICASSP39728.2021.9413392>

AUTHORS: D. E. O. Tzamaras, E. Pavez, B. Girault, A. Ortega, I. Blanes and J. Serra-Sagrista

HIGHLIGHT: By exploiting a new extension of the admissibility condition in GraphBior we propose a new fundamental matrix with the goal of distributing the errors of GraphBior more uniformly across pixels with different node degrees.

1104, TITLE: Graph Signal Compression via Task-Based Quantization

<https://doi.org/10.1109/ICASSP39728.2021.9414657>

AUTHORS: P. Li, N. Shlezinger, H. Zhang, B. Wang and Y. C. Eldar

HIGHLIGHT: In this work we study the joint design of graph signal sampling along with the quantization of these samples, for graph signal compression.

1105, TITLE: A Partially Collapsed Gibbs Sampler for Unsupervised Nonnegative Sparse Signal Restoration

<https://doi.org/10.1109/ICASSP39728.2021.9414293>

AUTHORS: M. C. Amrouche, H. Carfantan and J. Idier

HIGHLIGHT: In this paper the problem of restoration of unsupervised nonnegative sparse signals is addressed in the Bayesian framework.

1106, TITLE: A Structure-Guided and Sparse-Representation-Based 3d Seismic Inversion Method

<https://doi.org/10.1109/ICASSP39728.2021.9415071>

AUTHORS: B. She, Y. Wang and G. Hu

HIGHLIGHT: In this work, with the help of gradient structure tensor (GST) and dictionary learning and sparse representation (DLSR) technologies, we propose a 3D inversion approach (GST-DLSR) that considers both vertical and horizontal structural constraints.

1107, TITLE: Accelerating Frank-Wolfe with Weighted Average Gradients

<https://doi.org/10.1109/ICASSP39728.2021.9414485>

AUTHORS: Y. Zhang, B. Li and G. B. Giannakis

HIGHLIGHT: The present contribution broadens its scope by replacing the gradient per FW subproblem with a weighted average of gradients.

- 1108, TITLE: Yapa: Accelerated Proximal Algorithm for Convex Composite Problems  
<https://doi.org/10.1109/ICASSP39728.2021.9414656>  
AUTHORS: G. Chierchia and M. El Gheche  
HIGHLIGHT: In this paper, we propose a primal proximal method derived from a three-operator splitting in a product space and accelerated with Anderson extrapolation.
- 1109, TITLE: Data Discovery Using Lossless Compression-Based Sparse Representation  
<https://doi.org/10.1109/ICASSP39728.2021.9415027>  
AUTHORS: E. Sabeti, P. X. K. Song and A. O. Hero  
HIGHLIGHT: In this paper, we propose a data-driven sparse representation using orthonormal bases under the lossless compression constraint.
- 1110, TITLE: Safe Screening for Sparse Regression with the Kullback-Leibler Divergence  
<https://doi.org/10.1109/ICASSP39728.2021.9414183>  
AUTHORS: C. F. Dantas, E. Soubies and C. Fvotte  
HIGHLIGHT: In this paper, we extend the GAP Safe screening rule to the l1-regularized Kullback-Leibler divergence which does not fulfil the regularity assumptions made in previous works.
- 1111, TITLE: On the Convergence of Randomized Bregman Coordinate Descent for Non-Lipschitz Composite Problems  
<https://doi.org/10.1109/ICASSP39728.2021.9414191>  
AUTHORS: T. Gao, S. Lu, J. Liu and C. Chu  
HIGHLIGHT: We propose a new randomized Bregman (block) coordinate descent (RBCD) method for minimizing a composite problem, where the objective function could be either convex or nonconvex, and the smooth part are freed from the global Lipschitz-continuous (partial) gradient assumption.
- 1112, TITLE: A Global Cayley Parametrization of Stiefel Manifold for Direct Utilization of Optimization Mechanisms Over Vector Spaces  
<https://doi.org/10.1109/ICASSP39728.2021.9414157>  
AUTHORS: K. Kume and I. Yamada  
HIGHLIGHT: In this paper, we present a global parametrization of the Stiefel manifold entirely by a single fixed vector space with the Cayley transform, say Global Cayley Parametrization (G-CP), to solve the problem through optimization over a vector space.
- 1113, TITLE: Training Logical Neural Networks by Primal?Dual Methods for Neuro-Symbolic Reasoning  
<https://doi.org/10.1109/ICASSP39728.2021.9415044>  
AUTHORS: S. Lu, N. Khan, I. Y. Akhalwaya, R. Riegel, L. Horesh and A. Gray  
HIGHLIGHT: In this paper, we propose a unified framework for solving this nonlinear programming problem by leveraging primal-dual optimization methods, and quantify the corresponding convergence rate to the Karush-Kuhn-Tucker (KKT) points of this problem.
- 1114, TITLE: Cooperative Parameter Tracking on the Unit Sphere Using Distributed Adapt-Then-Combine Particle Filters and Parallel Transport  
<https://doi.org/10.1109/ICASSP39728.2021.9414948>  
AUTHORS: C. G. de Figueredo, C. J. Bordin and M. G. S. Bruno  
HIGHLIGHT: This paper introduces a new distributed Adapt-then-Combine (ATC) diffusion algorithm for cooperative tracking of an un-known state vector that evolves on the unit hypersphere.
- 1115, TITLE: Bayesian Estimation of a Tail-Index with Marginalized Threshold  
<https://doi.org/10.1109/ICASSP39728.2021.9413935>  
AUTHORS: D. E. Johnston and P. M. Djuric  
HIGHLIGHT: In this paper, we develop a new method for estimating the tail-index found in extreme value statistics.
- 1116, TITLE: Block Kalman Filter: An Asymptotic Block Particle Filter in the Linear Gaussian Case  
<https://doi.org/10.1109/ICASSP39728.2021.9413963>  
AUTHORS: R. Min, C. Garnier, F. Septier and J. Klein  
HIGHLIGHT: In this paper, we introduce the block based approach in the Kalman filter and show that the block particle filter asymptotically acts as the resulting block Kalman filter.
- 1117, TITLE: Particle Gibbs Sampling for Regime-Switching State-Space Models  
<https://doi.org/10.1109/ICASSP39728.2021.9414875>  
AUTHORS: Y. El-Laham, L. Yang, H. J. Lynch, P. M. Djuric and M. F. Bugallo  
HIGHLIGHT: In this work, we develop a particle Gibbs sampling algorithm for Bayesian learning in RS-SSMs.



- 1118, TITLE: Adaptive Importance Sampling Via Auto-Regressive Generative Models and Gaussian Processes  
<https://doi.org/10.1109/ICASSP39728.2021.9414734>  
AUTHORS: H. Wang, M. F. Bugallo and P. M. Djuric  
HIGHLIGHT: In this paper, we introduce a class of adaptive importance sampling methods where the proposal distribution is constructed in a way that Gaussian processes are combined autoregressively.
- 1119, TITLE: Variational Parameter Learning in Sequential State-Space Model Via Particle Filtering  
<https://doi.org/10.1109/ICASSP39728.2021.9414033>  
AUTHORS: C. Li and S. J. Godsill  
HIGHLIGHT: We present a novel algorithm, the particle filter variational inference (PF-VI) algorithm, which achieves closed-form learning of SSM parameters while tractably inferring the non-linear sequential states.
- 1120, TITLE: Correlation-Based Robust Linear Regression with Iterative Outlier Removal  
<https://doi.org/10.1109/ICASSP39728.2021.9414849>  
AUTHORS: J. Ding, J. Wang, Y. Zhang, Y. Li and N. Zheng  
HIGHLIGHT: Here we consider linear regression from the view of correlation and propose a robust regression algorithm.
- 1121, TITLE: On the Optimality of Backward Regression: Sparse Recovery and Subset Selection  
<https://doi.org/10.1109/ICASSP39728.2021.9415082>  
AUTHORS: S. Ament and C. Gomes  
HIGHLIGHT: We present novel guarantees for the algorithm, propose an efficient, numerically stable implementation, and put forth Stepwise Regression with Replacement (SRR), a new family of two-stage algorithms that employs both forward and backward steps for compressed sensing problems.
- 1122, TITLE: General Total Variation Regularized Sparse Bayesian Learning for Robust Block-Sparse Signal Recovery  
<https://doi.org/10.1109/ICASSP39728.2021.9413977>  
AUTHORS: A. Sant, M. Leinonen and B. D. Rao  
HIGHLIGHT: We propose a novel Sparse Bayesian Learning (SBL) method for block-sparse recovery based on popular CS based regularizers with the function input variable related to total variation (TV).
- 1123, TITLE: Automatic Registration and Clustering of Time Series  
<https://doi.org/10.1109/ICASSP39728.2021.9414417>  
AUTHORS: M. Weylandt and G. Michailidis  
HIGHLIGHT: We propose a new method for automatic time series alignment within a clustering problem.
- 1124, TITLE: Low-Rank on Graphs Plus Temporally Smooth Sparse Decomposition for Anomaly Detection in Spatiotemporal Data  
<https://doi.org/10.1109/ICASSP39728.2021.9414360>  
AUTHORS: S. E. Sofuoglu and S. Aviyente  
HIGHLIGHT: We introduce an unsupervised tensor-based anomaly detection method that takes the sparse and temporally continuous nature of anomalies into account.
- 1125, TITLE: A Parallel Algorithm for Phase Retrieval with Dictionary Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9413991>  
AUTHORS: T. Liu, A. M. Tillmann, Y. Yang, Y. C. Eldar and M. Pesavento  
HIGHLIGHT: We propose a new formulation for the joint phase retrieval and dictionary learning problem with a reduced number of regularization parameters to be tuned.
- 1126, TITLE: Improving RNN Transducer Modeling for Small-Footprint Keyword Spotting  
<https://doi.org/10.1109/ICASSP39728.2021.9414339>  
AUTHORS: Y. Tian, H. Yao, M. Cai, Y. Liu and Z. Ma  
HIGHLIGHT: In this paper, we improve the RNN-T modeling for small-footprint keyword spotting in three aspects.
- 1127, TITLE: Cascaded Encoders for Unifying Streaming and Non-Streaming ASR  
<https://doi.org/10.1109/ICASSP39728.2021.9414607>  
AUTHORS: A. Narayanan et al.  
HIGHLIGHT: This work presents cascaded encoders for building a single E2E ASR model that can operate in both these modes simultaneously.

- 1128, TITLE: A Better and Faster end-to-end Model for Streaming ASR  
<https://doi.org/10.1109/ICASSP39728.2021.9413899>  
AUTHORS: B. Li et al.  
HIGHLIGHT: To address this issue, we look at encouraging the E2E model to emit words early, through an algorithm called FastEmit [3].
- 1129, TITLE: Efficient Knowledge Distillation for RNN-Transducer Models  
<https://doi.org/10.1109/ICASSP39728.2021.9413905>  
AUTHORS: S. Panchapagesan, D. S. Park, C. -C. Chiu, Y. Shangguan, Q. Liang and A. Gruenstein  
HIGHLIGHT: In this paper, we develop a distillation method for RNN-Transducer (RNN-T) models, a popular end-to-end neural network architecture for streaming speech recognition.
- 1130, TITLE: Phoneme Based Neural Transducer for Large Vocabulary Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9413648>  
AUTHORS: W. Zhou, S. Berger, R. Schl?ter and H. Ney  
HIGHLIGHT: To join the advantages of classical and end-to-end approaches for speech recognition, we present a simple, novel and competitive approach for phoneme-based neural transducer modeling.
- 1131, TITLE: RNN-T Based Open-Vocabulary Keyword Spotting in Mandarin with Multi-Level Detection  
<https://doi.org/10.1109/ICASSP39728.2021.9413588>  
AUTHORS: Z. Liu, T. Li and P. Zhang  
HIGHLIGHT: In this paper, we propose an RNN Transducer (RNN-T) based keyword spotting system with a constrained attention mechanism biasing module that biases the RNN-T model towards a specific keyword of interest.
- 1132, TITLE: Advancing RNN Transducer Technology for Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9414716>  
AUTHORS: G. Saon, Z. T'ske, D. Bolanos and B. Kingsbury  
HIGHLIGHT: We investigate a set of techniques for RNN Transducers (RNN-Ts) that were instrumental in lowering the word error rate on three different tasks (Switchboard 300 hours, conversational Spanish 780 hours and conversational Italian 900 hours).
- 1133, TITLE: Less is More: Improved RNN-T Decoding Using Limited Label Context and Path Merging  
<https://doi.org/10.1109/ICASSP39728.2021.9414212>  
AUTHORS: R. Prabhavalkar et al.  
HIGHLIGHT: In this work, we study the influence of the amount of label context on the model's accuracy, and its impact on the efficiency of the decoding process.
- 1134, TITLE: Simpleflat: A Simple Whole-Network Pre-Training Approach for RNN Transducer-Based End-to-End Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9413741>  
AUTHORS: T. Moriya et al.  
HIGHLIGHT: Our solution is SimpleFlat (SF), a novel and simple whole-network pretraining approach for RNN-T.
- 1135, TITLE: Echo State Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9414495>  
AUTHORS: H. Shrivastava, A. Garg, Y. Cao, Y. Zhang and T. Sainath  
HIGHLIGHT: We propose automatic speech recognition (ASR) models inspired by echo state network (ESN) [1], in which a subset of recurrent neural networks (RNN) layers in the models are randomly initialized and untrained.
- 1136, TITLE: Using Synthetic Audio to Improve the Recognition of Out-of-Vocabulary Words in End-to-End Asr Systems  
<https://doi.org/10.1109/ICASSP39728.2021.9414778>  
AUTHORS: X. Zheng, Y. Liu, D. Gunceler and D. Willett  
HIGHLIGHT: In this paper, we use a text-to-speech (TTS) engine to provide synthetic audio for out-of-vocabulary (OOV) words.
- 1137, TITLE: Wave-Tacotron: Spectrogram-Free End-to-End Text-to-Speech Synthesis  
<https://doi.org/10.1109/ICASSP39728.2021.9413851>  
AUTHORS: R. J. Weiss, R. Skerry-Ryan, E. Battenberg, S. Mariooryad and D. P. Kingma  
HIGHLIGHT: We describe a sequence-to-sequence neural network which directly generates speech waveforms from text inputs.

- 1138, TITLE: Patnet : A Phoneme-Level Autoregressive Transformer Network for Speech Synthesis  
<https://doi.org/10.1109/ICASSP39728.2021.9413658>  
AUTHORS: S. Wang, Z. Ling, R. Fu, J. Yi and J. Tao  
HIGHLIGHT: Aiming at efficiently predicting acoustic features with high naturalness and robustness, this paper proposes PATNet, a neural acoustic model for speech synthesis using phoneme-level autoregression.
- 1139, TITLE: Multi-Rate Attention Architecture for Fast Streamable Text-to-Speech Spectrum Modeling  
<https://doi.org/10.1109/ICASSP39728.2021.9414809>  
AUTHORS: Q. He, Z. Xiu, T. Koehler and J. Wu  
HIGHLIGHT: In this paper, we propose a multi-rate attention architecture that breaks the latency and RTF bottlenecks by computing a compact representation during encoding and recurrently generating the attention vector in a streaming manner during decoding.
- 1140, TITLE: End-to-End Text-to-Speech Using Latent Duration Based on VQ-VAE  
<https://doi.org/10.1109/ICASSP39728.2021.9414499>  
AUTHORS: Y. Yasuda, X. Wang and J. Yamagishid  
HIGHLIGHT: We propose a new TTS framework using explicit duration modeling that incorporates duration as a discrete latent variable to TTS and enables joint optimization of whole modules from scratch.
- 1141, TITLE: Lightspeech: Lightweight and Fast Text to Speech with Neural Architecture Search  
<https://doi.org/10.1109/ICASSP39728.2021.9414403>  
AUTHORS: R. Luo et al.  
HIGHLIGHT: In this paper, we propose LightSpeech, which leverages neural architecture search (NAS) to automatically design more lightweight and efficient models based on FastSpeech.
- 1142, TITLE: A New High Quality Trajectory Tiling Based Hybrid TTS In Real Time  
<https://doi.org/10.1109/ICASSP39728.2021.9413969>  
AUTHORS: F. -L. Xie, X. -H. Li, W. -C. Su, L. Lu and F. K. Soong  
HIGHLIGHT: A trajectory tiling based, hybrid TTS is revisited in this study for improving its synthesis performance.
- 1143, TITLE: Parallel Tacotron: Non-Autoregressive and Controllable TTS  
<https://doi.org/10.1109/ICASSP39728.2021.9414718>  
AUTHORS: I. Elias et al.  
HIGHLIGHT: This paper proposes a non-autoregressive neural text-to-speech model augmented with a variational autoencoder-based residual encoder.
- 1144, TITLE: Fcl-Taco2: Towards Fast, Controllable and Lightweight Text-to-Speech Synthesis  
<https://doi.org/10.1109/ICASSP39728.2021.9414870>  
AUTHORS: D. Wang et al.  
HIGHLIGHT: To achieve fast inference speed and small model size while maintain high-quality speech, we propose FCL-taco2, a Fast, Controllable and Lightweight (FCL) TTS model based on Tacotron2.
- 1145, TITLE: Prosodic Clustering for Phoneme-Level Prosody Control in End-to-End Speech Synthesis  
<https://doi.org/10.1109/ICASSP39728.2021.9413604>  
AUTHORS: A. Vioni et al.  
HIGHLIGHT: This paper presents a method for controlling the prosody at the phoneme level in an autoregressive attention-based text-to-speech system.
- 1146, TITLE: Improving Naturalness and Controllability of Sequence-to-Sequence Speech Synthesis by Learning Local Prosody Representations  
<https://doi.org/10.1109/ICASSP39728.2021.9414720>  
AUTHORS: C. Gong, L. Wang, Z. Ling, S. Guo, J. Zhang and J. Dang  
HIGHLIGHT: In this study, we extended Tacotron2 with a pitch prediction task to capture discrete pitch-related representations.
- 1147, TITLE: Multi-Speaker Emotional Speech Synthesis with Fine-Grained Prosody Modeling  
<https://doi.org/10.1109/ICASSP39728.2021.9413398>  
AUTHORS: C. Lu, X. Wen, R. Liu and X. Chen  
HIGHLIGHT: We present an end-to-end system for multi-speaker emotional speech synthesis.

- 1148, TITLE: Emotion Controllable Speech Synthesis Using Emotion-Unlabeled Dataset with the Assistance of Cross-Domain Speech Emotion Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9413907>  
AUTHORS: X. Cai, D. Dai, Z. Wu, X. Li, J. Li and H. Meng  
HIGHLIGHT: In this paper, we propose a novel approach for emotional TTS synthesis on a TTS dataset without emotion labels.
- 1149, TITLE: Dual-Path Modeling for Long Recording Speech Separation in Meetings  
<https://doi.org/10.1109/ICASSP39728.2021.9414127>  
AUTHORS: C. Li et al.  
HIGHLIGHT: In this work, we further extend the dual-path modeling framework for CSS task. A transformer-based dual-path system is proposed, which integrates transform layers for global modeling.
- 1150, TITLE: Time-Domain Loss Modulation Based on Overlap Ratio for Monaural Conversational Speaker Separation  
<https://doi.org/10.1109/ICASSP39728.2021.9413900>  
AUTHORS: H. Taherian and D. Wang  
HIGHLIGHT: We propose a new loss function for speaker separation based on permutation invariant training that dynamically reweighs losses using the segment overlap ratio.
- 1151, TITLE: Continuous Speech Separation with Conformer  
<https://doi.org/10.1109/ICASSP39728.2021.9413423>  
AUTHORS: S. Chen et al.  
HIGHLIGHT: This paper examines the use of Conformer architecture in lieu of recurrent neural networks for the separation model.
- 1152, TITLE: A Flow-Based Neural Network for Time Domain Speech Enhancement  
<https://doi.org/10.1109/ICASSP39728.2021.9413999>  
AUTHORS: M. Strauss and B. Edler  
HIGHLIGHT: Thus, in this paper we propose a NF framework to directly model the enhancement process by density estimation of clean speech utterances conditioned on their noisy counterpart.
- 1153, TITLE: Sandglassnet: A Light Multi-Granularity Self-Attentive Network for Time-Domain Speech Separation  
<https://doi.org/10.1109/ICASSP39728.2021.9413837>  
AUTHORS: M. W. Y. Lam, J. Wang, D. Su and D. Yu  
HIGHLIGHT: We introduce a self-attentive network with a novel sandglass-shape, namely Sandglassnet, which advances the state-of-the-art (SOTA) SS performance at significantly smaller model size and computational cost.
- 1154, TITLE: TransMask: A Compact and Fast Speech Separation Model Based on Transformer  
<https://doi.org/10.1109/ICASSP39728.2021.9413670>  
AUTHORS: Z. Zhang, B. He and Z. Zhang  
HIGHLIGHT: To make these models more practical by reducing the model size and inference time while maintaining high separation quality, we propose a new transformer-based speech separation approach, called TransMask.
- 1155, TITLE: One Shot Learning for Speech Separation  
<https://doi.org/10.1109/ICASSP39728.2021.9413956>  
AUTHORS: Y. -K. Wu, K. -P. Huang, Y. Tsao and H. -y. Lee  
HIGHLIGHT: In this paper, we use model-agnostic meta-learning(MAML) algorithm and almost no inner loop(ANIL) algorithm in Conv-TasNet to achieve this goal.
- 1156, TITLE: Training Noisy Single-Channel Speech Separation with Noisy Oracle Sources: A Large Gap and a Small Step  
<https://doi.org/10.1109/ICASSP39728.2021.9413975>  
AUTHORS: M. Maciejewski, J. Shi, S. Watanabe and S. Khudanpur  
HIGHLIGHT: We demonstrate the relative inseparability of noise and that this noisy speech paradigm leads to significant degradation of system performance.
- 1157, TITLE: Speaker and Direction Inferred Dual-Channel Speech Separation  
<https://doi.org/10.1109/ICASSP39728.2021.9413818>  
AUTHORS: C. Li, J. Xu, N. Mesgarani and B. Xu  
HIGHLIGHT: In this work, we employ ideas from auditory attention with two ears and propose a speaker and direction inferred speech separation network (dubbed SDNet) to solve the cocktail party problem.

- 1158, TITLE: Speech Dereverberation Using Variational Autoencoders  
<https://doi.org/10.1109/ICASSP39728.2021.9414736>  
AUTHORS: D. Baby and H. Bourlard  
HIGHLIGHT: This paper presents a statistical method for single-channel speech dereverberation using a variational autoencoder (VAE) for modelling the speech spectra.
- 1159, TITLE: Real-Time Denoising and Dereverberation with Tiny Recurrent U-Net  
<https://doi.org/10.1109/ICASSP39728.2021.9414852>  
AUTHORS: H. -S. Choi, S. Park, J. H. Lee, H. Heo, D. Jeon and K. Lee  
HIGHLIGHT: To this end, we propose Tiny Recurrent U-Net (TRU-Net), a lightweight online inference model that matches the performance of current state-of-the-art models.
- 1160, TITLE: Weighted Magnitude-Phase Loss for Speech Dereverberation  
<https://doi.org/10.1109/ICASSP39728.2021.9414929>  
AUTHORS: J. Zhang, M. D. Plumbley and W. Wang  
HIGHLIGHT: We propose a new weighted magnitude-phase loss function, which is divided into a magnitude component and a phase component, to train a neural network to estimate complex ideal ratio masks.
- 1161, TITLE: Speaker Embeddings for Diarization of Broadcast Data In The Allies Challenge  
<https://doi.org/10.1109/ICASSP39728.2021.9414215>  
AUTHORS: A. Larcher et al.  
HIGHLIGHT: The work reported in this paper compares the performance of different embeddings extracted from MFCCs or the raw signal for speaker diarization and broadcast media treated with compression and sub-sampling, operations which typically degrade performance.
- 1162, TITLE: On the Detection of Pitch-Shifted Voice: Machines and Human Listeners  
<https://doi.org/10.1109/ICASSP39728.2021.9414890>  
AUTHORS: D. Looney and N. D. Gaubitch  
HIGHLIGHT: We present a performance comparison between human listeners and a simple algorithm for the task of speech anomaly detection.
- 1163, TITLE: The ins and outs of speaker recognition: lessons from VoxSRC 2020  
<https://doi.org/10.1109/ICASSP39728.2021.9413948>  
AUTHORS: Y. Kwon, H. -S. Heo, B. -J. Lee and J. S. Chung  
HIGHLIGHT: The goal of this work is robust speaker recognition of utterances recorded in these challenging environments.
- 1164, TITLE: The Idlab Voxsrc-20 Submission: Large Margin Fine-Tuning and Quality-Aware Score Calibration in DNN Based Speaker Verification  
<https://doi.org/10.1109/ICASSP39728.2021.9414600>  
AUTHORS: J. Thienpondt, B. Desplanques and K. Demuynck  
HIGHLIGHT: In this paper we propose and analyse a large margin fine-tuning strategy and a quality-aware score calibration in text-independent speaker verification.
- 1165, TITLE: Analysis of the but Diarization System for Voxconverse Challenge  
<https://doi.org/10.1109/ICASSP39728.2021.9414315>  
AUTHORS: F. Landini et al.  
HIGHLIGHT: This paper describes the system developed by the BUT team for the fourth track of the VoxCeleb Speaker Recognition Challenge, focusing on diarization on the VoxConverse dataset.
- 1166, TITLE: Microsoft Speaker Diarization System for the Voxceleb Speaker Recognition Challenge 2020  
<https://doi.org/10.1109/ICASSP39728.2021.9413832>  
AUTHORS: X. Xiao et al.  
HIGHLIGHT: This paper describes the Microsoft speaker diarization system for monaural multi-talker recordings in the wild, evaluated at the diarization track of the VoxCeleb Speaker Recognition Challenge (VoxSRC) 2020.
- 1167, TITLE: Squeezing Value of Cross-Domain Labels: A Decoupled Scoring Approach for Speaker Verification  
<https://doi.org/10.1109/ICASSP39728.2021.9414794>  
AUTHORS: L. Li, Y. Zhang, J. Kang, T. F. Zheng and D. Wang

**HIGHLIGHT:** Based on this analysis, we present a decoupled scoring approach that can maximally squeeze the value of cross-domain labels and obtain optimal verification scores in the enrollment-test mismatch condition.

1168, **TITLE:** Self-Supervised Learning Based Domain Adaptation for Robust Speaker Verification

<https://doi.org/10.1109/ICASSP39728.2021.9414261>

**AUTHORS:** Z. Chen, S. Wang and Y. Qian

**HIGHLIGHT:** In this paper, we incorporate self-supervised learning strategy to the unsupervised domain adaptation system and proposed a self-supervised learning based domain adaptation approach (SSDA).

1169, **TITLE:** Meta-Learning for Cross-Channel Speaker Verification

<https://doi.org/10.1109/ICASSP39728.2021.9413978>

**AUTHORS:** H. Zhang, L. Wang, K. A. Lee, M. Liu, J. Dang and H. Chen

**HIGHLIGHT:** To this end, we propose a meta speaker embedding network (MSEN) via meta-learning to generate channel-invariant utterance embeddings.

1170, **TITLE:** SynAug: Synthesis-Based Data Augmentation for Text-Dependent Speaker Verification

<https://doi.org/10.1109/ICASSP39728.2021.9414438>

**AUTHORS:** C. Du, B. Han, S. Wang, Y. Qian and K. Yu

**HIGHLIGHT:** In this work, we propose a synthesis based data augmentation method (SynAug) to expand the training set with more speakers and text-controlled synthesized speech.

1171, **TITLE:** Unit Selection Synthesis Based Data Augmentation for Fixed Phrase Speaker Verification

<https://doi.org/10.1109/ICASSP39728.2021.9414550>

**AUTHORS:** H. Huang, X. Xiang, F. Zhao, S. Wang and Y. Qian

**HIGHLIGHT:** In this work, we propose a unit selection synthesis based data augmentation method to leverage the abundant text-independent data resources.

1172, **TITLE:** Improving Speaker Verification in Reverberant Environments

<https://doi.org/10.1109/ICASSP39728.2021.9413731>

**AUTHORS:** X. Chen and S. A. Zahorian

**HIGHLIGHT:** In this paper, we first introduce a new feature set that gives more details in the frequency dimension in the 2-D time-frequency space used to represent speech.

1173, **TITLE:** Transformer-Transducers for Code-Switched Speech Recognition

<https://doi.org/10.1109/ICASSP39728.2021.9413562>

**AUTHORS:** S. Dalmia, Y. Liu, S. Ronanki and K. Kirchhoff

**HIGHLIGHT:** In this paper, we present an end-to-end ASR system using a transformer-transducer model architecture for code-switched speech recognition.

1174, **TITLE:** Wake Word Detection with Streaming Transformers

<https://doi.org/10.1109/ICASSP39728.2021.9414777>

**AUTHORS:** Y. Wang, H. Lv, D. Povey, L. Xie and S. Khudanpur

**HIGHLIGHT:** In this paper we explore the performance of several variants of chunk-wise streaming Transformers tailored for wake word detection in a recently proposed LF-MMI system, including looking-ahead to the next chunk, gradient stopping, different positional embedding methods and adding same-layer dependency between chunks.

1175, **TITLE:** Capturing Multi-Resolution Context by Dilated Self-Attention

<https://doi.org/10.1109/ICASSP39728.2021.9415001>

**AUTHORS:** N. Moritz, T. Hori and J. Le Roux

**HIGHLIGHT:** In this work, we propose a combination of restricted self-attention and a dilation mechanism, which we refer to as dilated self-attention.

1176, **TITLE:** Recent Developments on Espnet Toolkit Boosted By Conformer

<https://doi.org/10.1109/ICASSP39728.2021.9414858>

**AUTHORS:** P. Guo et al.

**HIGHLIGHT:** In this study, we present recent developments on ESPnet: End-to- End Speech Processing toolkit, which mainly involves a recently proposed architecture called Conformer, Convolution-augmented Transformer.

1177, **TITLE:** Hierarchical Transformer-Based Large-Context End-To-End ASR with Large-Context Knowledge Distillation

<https://doi.org/10.1109/ICASSP39728.2021.9414928>

AUTHORS: R. Masumura, N. Makishima, M. Ihori, A. Takashima, T. Tanaka and S. Orihashi  
HIGHLIGHT: We present a novel large-context end-to-end automatic speech recognition (E2E-ASR) model and its effective training method based on knowledge distillation.

1178, TITLE: End-to-End Multi-Channel Transformer for Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9414123>  
AUTHORS: F. -J. Chang, M. Radfar, A. Mouchtaris, B. King and S. Kunzmann  
HIGHLIGHT: In this paper, we leverage the neural transformer architectures for multi-channel speech recognition systems, where the spectral and spatial information collected from different microphones are integrated using attention layers.

1179, TITLE: CASS-NAT: CTC Alignment-Based Single Step Non-Autoregressive Transformer for Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9413429>  
AUTHORS: R. Fan, W. Chu, P. Chang and J. Xiao  
HIGHLIGHT: We propose a CTC alignment-based single step non-autoregressive transformer (CASS-NAT) for speech recognition.

1180, TITLE: Non-Autoregressive Transformer ASR with CTC-Enhanced Decoder Input  
<https://doi.org/10.1109/ICASSP39728.2021.9414694>  
AUTHORS: X. Song, Z. Wu, Y. Huang, C. Weng, D. Su and H. Meng  
HIGHLIGHT: To address this problem, we propose a CTC-enhanced NAR transformer, which generates target sequence by refining predictions of the CTC module.

1181, TITLE: Transformer-Based End-to-End Speech Recognition with Local Dense Synthesizer Attention  
<https://doi.org/10.1109/ICASSP39728.2021.9414353>  
AUTHORS: M. Xu, S. Li and X. -L. Zhang  
HIGHLIGHT: Motivated by the fact that dense synthesizer attention (DSA), which dispenses with dot products and pairwise interactions, achieved competitive results in many language processing tasks, in this paper, we first propose a DSA-based speech recognition, as an alternative to SA.

1182, TITLE: Developing Real-Time Streaming Transformer Transducer for Speech Recognition on Large-Scale Dataset  
<https://doi.org/10.1109/ICASSP39728.2021.9413535>  
AUTHORS: X. Chen, Y. Wu, Z. Wang, S. Liu and J. Li  
HIGHLIGHT: In this work, we explored the potential of Transformer Transducer (T-T) models for the first pass decoding with low latency and fast speed on a large-scale dataset.

1183, TITLE: Head-Synchronous Decoding for Transformer-Based Streaming ASR  
<https://doi.org/10.1109/ICASSP39728.2021.9414103>  
AUTHORS: M. Li, C. Zorila and R. Doddipatla  
HIGHLIGHT: To overcome these issues, here we propose a head-synchronous (HS) version of the DACS algorithm, where the boundary of attention is jointly detected by all the DACS heads in each decoder layer.

1184, TITLE: History Utterance Embedding Transformer LM for Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9414575>  
AUTHORS: K. Deng, G. Cheng, H. Miao, P. Zhang and Y. Yan  
HIGHLIGHT: In this paper, we propose the history utterance embedding Transformer LM (HTLM), which includes an embedding generation network for extracting contextual information contained in the history utterances and a main Transformer LM for current prediction.

1185, TITLE: Maskcyclegan-VC: Learning Non-Parallel Voice Conversion with Filling in Frames  
<https://doi.org/10.1109/ICASSP39728.2021.9414851>  
AUTHORS: T. Kaneko, H. Kameoka, K. Tanaka and N. Hojo  
HIGHLIGHT: As an alternative, we propose MaskCycleGAN-VC, which is another extension of CycleGAN-VC2 and is trained using a novel auxiliary task called filling in frames (FIF).

1186, TITLE: Non-Parallel Many-To-Many Voice Conversion by Knowledge Transfer from a Text-To-Speech Model  
<https://doi.org/10.1109/ICASSP39728.2021.9414757>  
AUTHORS: X. YU and B. Mak  
HIGHLIGHT: In this paper, we present a simple but novel framework to train a nonparallel many-to-many voice conversion (VC) model based on the encoder-decoder architecture.



- 1187, TITLE: Non-Parallel Many-To-Many Voice Conversion Using Local Linguistic Tokens  
<https://doi.org/10.1109/ICASSP39728.2021.9413540>  
AUTHORS: C. Wang and Y. Yu  
HIGHLIGHT: To address this issue, we propose the Local Linguistic Tokens (LLTs) model to learn high-quality speaker-invariant linguistic embeddings using the multi-head attention module, which has shown great success in extracting speaking style embeddings in Global Style Tokens (GSTs).
- 1188, TITLE: Crank: An Open-Source Software for Nonparallel Voice Conversion Based on Vector-Quantized Variational Autoencoder  
<https://doi.org/10.1109/ICASSP39728.2021.9413959>  
AUTHORS: K. Kobayashi, W. -C. Huang, Y. -C. Wu, P. L. Tobing, T. Hayashi and T. Toda  
HIGHLIGHT: In this paper, we present an open-source software for developing a nonparallel voice conversion (VC) system named crank.
- 1189, TITLE: Fragmentvc: Any-To-Any Voice Conversion by End-To-End Extracting and Fusing Fine-Grained Voice Fragments with Attention  
<https://doi.org/10.1109/ICASSP39728.2021.9413699>  
AUTHORS: Y. Y. Lin, C. -M. Chien, J. -H. Lin, H. -y. Lee and L. -s. Lee  
HIGHLIGHT: In this paper we proposed FragmentVC, in which the latent phonetic structure of the utterance from the source speaker is obtained from Wav2Vec 2.0, while the spectral features of the utterance(s) from the target speaker are obtained from log mel-spectrograms.
- 1190, TITLE: Any-to-One Sequence-to-Sequence Voice Conversion Using Self-Supervised Discrete Speech Representations  
<https://doi.org/10.1109/ICASSP39728.2021.9415079>  
AUTHORS: W. -C. Huang, Y. -C. Wu and T. Hayashi  
HIGHLIGHT: We present a novel approach to any-to-one (A2O) voice conversion (VC) in a sequence-to-sequence (seq2seq) framework.
- 1191, TITLE: Towards Low-Resource Stargan Voice Conversion Using Weight Adaptive Instance Normalization  
<https://doi.org/10.1109/ICASSP39728.2021.9415042>  
AUTHORS: M. Chen, Y. Shi and T. Hain  
HIGHLIGHT: In this work, we aim at improving the data efficiency of the model and achieving a many-to-many non-parallel StarGAN-based voice conversion for a relatively large number of speakers with limited training samples.
- 1192, TITLE: Again-VC: A One-Shot Voice Conversion Using Activation Guidance and Adaptive Instance Normalization  
<https://doi.org/10.1109/ICASSP39728.2021.9414257>  
AUTHORS: Y. -H. Chen, D. -Y. Wu, T. -H. Wu and H. -y. Lee  
HIGHLIGHT: In this work, we propose AGAIN-VC, an innovative VC system using Activation Guidance and Adaptive Instance Normalization.
- 1193, TITLE: One-Shot Voice Conversion Based on Speaker Aware Module  
<https://doi.org/10.1109/ICASSP39728.2021.9414081>  
AUTHORS: Y. Zhang, H. Che, J. Li, C. Li, X. Wang and Z. Wang  
HIGHLIGHT: In this paper, we propose a speaker-aware voice conversion (SAVC) system realizing one-shot voice conversion without an adaptation stage.
- 1194, TITLE: Zero-Shot Voice Conversion with Adjusted Speaker Embeddings and Simple Acoustic Features  
<https://doi.org/10.1109/ICASSP39728.2021.9414975>  
AUTHORS: Z. Tan, J. Wei, J. Xu, Y. He and W. Lu  
HIGHLIGHT: In our work, a newly designed neural network was used to adjust the speaker embeddings of unseen speakers.
- 1195, TITLE: Towards Natural and Controllable Cross-Lingual Voice Conversion Based on Neural TTS Model and Phonetic Posteriorgram  
<https://doi.org/10.1109/ICASSP39728.2021.9414788>  
AUTHORS: S. Zhao, H. Wang, T. H. Nguyen and B. Ma  
HIGHLIGHT: In this paper, we build upon the neural text-to-speech (TTS) model, i.e., FastSpeech, and LPCNet neural vocoder to design a new cross-lingual VC framework named FastSpeech-VC.
- 1196, TITLE: Meta-Learning for Improving Rare Word Recognition in End-to-End ASR  
<https://doi.org/10.1109/ICASSP39728.2021.9414298>  
AUTHORS: F. Lux and N. T. Vu

**HIGHLIGHT:** In this work we take on the challenge of rare word recognition in end-to-end (E2E) automatic speech recognition (ASR) by integrating a meta learning mechanism into an E2E ASR system, enabling few-shot adaptation.

1197, **TITLE:** A Comparison of Methods for OOV-Word Recognition on a New Public Dataset

<https://doi.org/10.1109/ICASSP39728.2021.9415124>

**AUTHORS:** R. A. Braun, S. Madikeri and P. Motlicek

**HIGHLIGHT:** We showcase very large improvements in OOV-word recognition and make both the data and code available.

1198, **TITLE:** Convolutional Dropout and Wordpiece Augmentation for End-to-End Speech Recognition

<https://doi.org/10.1109/ICASSP39728.2021.9415004>

**AUTHORS:** H. Xu, Y. Huang, Y. Zhu, K. Audhkhasi and B. Ramabhadran

**HIGHLIGHT:** We propose a generalization of dropout, called "convolutional dropout", where each neuron's activation is replaced with a randomly-weighted linear combination of neuron values in its neighborhood.

1199, **TITLE:** Partially Overlapped Inference for Long-Form Speech Recognition

<https://doi.org/10.1109/ICASSP39728.2021.9414941>

**AUTHORS:** T. G. Kang, H. -G. Kim, M. -J. Lee, J. Lee and H. Lee

**HIGHLIGHT:** In this paper, we propose a more effective way of overlapped inference by aligning partially matched hypotheses.

1200, **TITLE:** Focus on the Present: A Regularization Method for the ASR Source-Target Attention Layer

<https://doi.org/10.1109/ICASSP39728.2021.9414648>

**AUTHORS:** N. Chen, P. Zelasko, J. Villalba and N. Dehak

**HIGHLIGHT:** This paper introduces a novel method to diagnose the source-target attention in state-of-the-art end-to-end speech recognition models with joint connectionist temporal classification (CTC) and attention training.

1201, **TITLE:** Bifocal Neural ASR: Exploiting Keyword Spotting for Inference Optimization

<https://doi.org/10.1109/ICASSP39728.2021.9414652>

**AUTHORS:** J. Macoskey, G. P. Strimel and A. Rastrow

**HIGHLIGHT:** We present Bifocal RNN-T, a new variant of the Recurrent Neural Network Transducer (RNN-T) architecture designed for improved inference time latency on speech recognition tasks.

1202, **TITLE:** FastEmit: Low-Latency Streaming ASR with Sequence-Level Emission Regularization

<https://doi.org/10.1109/ICASSP39728.2021.9413803>

**AUTHORS:** J. Yu et al.

**HIGHLIGHT:** In this work, we propose a sequence-level emission regularization method, named FastEmit, that applies latency regularization directly on per-sequence probability in training transducer models, and does not require any alignment.

1203, **TITLE:** Sparsification via Compressed Sensing for Automatic Speech Recognition

<https://doi.org/10.1109/ICASSP39728.2021.9413844>

**AUTHORS:** K. Zhen, H. D. Nguyen, F. -J. Chang, A. Mouchtaris and A. Rastrow

**HIGHLIGHT:** In this work, we propose a compressed sensing based pruning (CSP) approach to effectively address those questions.

1204, **TITLE:** Dynamic Sparsity Neural Networks for Automatic Speech Recognition

<https://doi.org/10.1109/ICASSP39728.2021.9414505>

**AUTHORS:** Z. Wu, D. Zhao, Q. Liang, J. Yu, A. Gulati and R. Pang

**HIGHLIGHT:** In this paper, we present Dynamic Sparsity Neural Networks (DSNN) that, once trained, can instantly switch to any predefined sparsity configuration at run-time.

1205, **TITLE:** An Asynchronous WFST-Based Decoder for Automatic Speech Recognition

<https://doi.org/10.1109/ICASSP39728.2021.9414509>

**AUTHORS:** H. Lv, Z. Chen, H. Xu, D. Povey, L. Xie and S. Khudanpur

**HIGHLIGHT:** We introduce asynchronous dynamic decoder, which adopts an efficient A\* algorithm to incorporate big language models in the one-pass decoding for large vocabulary continuous speech recognition.

1206, **TITLE:** Tiny Transducer: A Highly-Efficient Speech Recognition Model on Edge Devices

<https://doi.org/10.1109/ICASSP39728.2021.9413854>

**AUTHORS:** Y. Zhang, S. Sun and L. Ma

**HIGHLIGHT:** This paper proposes an extremely lightweight phone-based transducer model with a tiny decoding graph on edge devices.

1207, **TITLE:** Noise Level Limited Sub-Modeling for Diffusion Probabilistic Vocoders  
<https://doi.org/10.1109/ICASSP39728.2021.9415087>

**AUTHORS:** T. Okamoto, T. Toda, Y. Shiga and H. Kawai

**HIGHLIGHT:** This paper proposes a simple but effective noise level-limited sub-modeling framework for diffusion probabilistic vocoders Sub-WaveGrad and Sub-DiffWave.

1208, **TITLE:** StyleMelGAN: An Efficient High-Fidelity Adversarial Vocoder with Temporal Adaptive Normalization  
<https://doi.org/10.1109/ICASSP39728.2021.9413605>

**AUTHORS:** A. Mustafa, N. Pia and G. Fuchs

**HIGHLIGHT:** We therefore propose StyleMelGAN, a lightweight neural vocoder allowing synthesis of high-fidelity speech with low computational complexity.

1209, **TITLE:** Parallel Waveform Synthesis Based on Generative Adversarial Networks with Voicing-Aware Conditional Discriminators  
<https://doi.org/10.1109/ICASSP39728.2021.9413369>

**AUTHORS:** R. Yamamoto, E. Song, M. -J. Hwang and J. -M. Kim

**HIGHLIGHT:** This paper proposes voicing-aware conditional discriminators for Parallel WaveGAN-based waveform synthesis systems.

1210, **TITLE:** Universal Neural Vocoding with Parallel Wavenet  
<https://doi.org/10.1109/ICASSP39728.2021.9414444>

**AUTHORS:** Y. Jiao, A. Gabrys, G. Tinchev, B. Putrycz, D. Korzekwa and V. Klimkov

**HIGHLIGHT:** We present a universal neural vocoder based on Parallel WaveNet, with an additional conditioning network called Audio Encoder.

1211, **TITLE:** Periodnet: A Non-Autoregressive Waveform Generation Model with a Structure Separating Periodic and Aperiodic Components  
<https://doi.org/10.1109/ICASSP39728.2021.9414401>

**AUTHORS:** Y. Hono, S. Takaki, K. Hashimoto, K. Oura, Y. Nankaku and K. Tokuda

**HIGHLIGHT:** We propose PeriodNet, a non-autoregressive (non-AR) waveform generation model with a new model structure for modeling periodic and aperiodic components in speech waveforms.

1212, **TITLE:** LVCNet: Efficient Condition-Dependent Modeling Network for Waveform Generation  
<https://doi.org/10.1109/ICASSP39728.2021.9414710>

**AUTHORS:** Z. Zeng, J. Wang, N. Cheng and J. Xiao

**HIGHLIGHT:** In this paper, we propose a novel conditional convolution network, named location-variable convolution, to model the dependencies of the waveform sequence.

1213, **TITLE:** Graphspeech: Syntax-Aware Graph Attention Network for Neural Speech Synthesis  
<https://doi.org/10.1109/ICASSP39728.2021.9413513>

**AUTHORS:** R. Liu, B. Sisman and H. Li

**HIGHLIGHT:** We propose a novel neural TTS model, denoted as GraphSpeech, that is formulated under graph neural network framework.

1214, **TITLE:** Syntactic Representation Learning For Neural Network Based TTS with Syntactic Parse Tree Traversal  
<https://doi.org/10.1109/ICASSP39728.2021.9414671>

**AUTHORS:** C. Song, J. Li, Y. Zhou, Z. Wu and H. Meng

**HIGHLIGHT:** In this paper, we propose a syntactic representation learning method based on syntactic parse tree traversal to automatically utilize the syntactic structure information.

1215, **TITLE:** A Chapter-Wise Understanding System for Text-To-Speech in Chinese Novels  
<https://doi.org/10.1109/ICASSP39728.2021.9415078>

**AUTHORS:** J. Pan, L. Wu, X. Yin, P. Wu, C. Xu and Z. Ma

**HIGHLIGHT:** In this paper, we propose a chapter-wise understanding system for Chinese novels, to predict speaker and emotion tags automatically based on the chapter-level context.

1216, **TITLE:** A Universal Bert-Based Front-End Model for Mandarin Text-To-Speech Synthesis

<https://doi.org/10.1109/ICASSP39728.2021.9414935>

AUTHORS: Z. Bai and B. Hu

HIGHLIGHT: In this paper, we propose a universal BERT-based model that can be used for various tasks in the Mandarin front-end without changing its architecture.

1217, TITLE: Improving Prosody Modelling with Cross-Utterance Bert Embeddings for End-to-End Speech Synthesis

<https://doi.org/10.1109/ICASSP39728.2021.9414102>

AUTHORS: G. Xu, W. Song, Z. Zhang, C. Zhang, X. He and B. Zhou

HIGHLIGHT: In this paper, we propose to use the text embeddings of the neighboring sentences to improve the prosody generation for each utterance of a paragraph in an end-to-end fashion without using any explicit prosody features.

1218, TITLE: Time-Domain Speech Extraction with Spatial Information and Multi Speaker Conditioning Mechanism

<https://doi.org/10.1109/ICASSP39728.2021.9414092>

AUTHORS: J. Zhang, C. Zorila, R. Doddipatla and J. Barker

HIGHLIGHT: In this paper, we present a novel multi-channel speech extraction system to simultaneously extract multiple clean individual sources from a mixture in noisy and reverberant environments.

1219, TITLE: ADL-MVDR: All Deep Learning MVDR Beamformer for Target Speech Separation

<https://doi.org/10.1109/ICASSP39728.2021.9413594>

AUTHORS: Z. Zhang, Y. Xu, M. Yu, S. -X. Zhang, L. Chen and D. Yu

HIGHLIGHT: In this paper, we propose a novel all deep learning MVDR framework, where the matrix inversion and eigenvalue decomposition are replaced by two recurrent neural networks (RNNs), to resolve both issues at the same time.

1220, TITLE: Multi-Channel Target Speech Extraction with Channel Decorrelation and Target Speaker Adaptation

<https://doi.org/10.1109/ICASSP39728.2021.9414244>

AUTHORS: J. Han, X. Zhou, Y. Long and Y. Li

HIGHLIGHT: In this work, we propose two methods for exploiting the multi-channel spatial information to extract the target speech.

1221, TITLE: Speaker Activity Driven Neural Speech Extraction

<https://doi.org/10.1109/ICASSP39728.2021.9414998>

AUTHORS: M. Delcroix, K. Zmolikova, T. Ochiai, K. Kinoshita and T. Nakatani

HIGHLIGHT: In this paper, we explore the use of speaker activity information as an auxiliary clue for single-channel neural network-based speech extraction.

1222, TITLE: Wase: Learning When to Attend for Speaker Extraction in Cocktail Party Environments

<https://doi.org/10.1109/ICASSP39728.2021.9413411>

AUTHORS: Y. Hao, J. Xu, P. Zhang and B. Xu

HIGHLIGHT: Inspired by it, we explicitly modeled the onset cue and verified the effectiveness in the speaker extraction task.

1223, TITLE: Multi-Stage Speaker Extraction with Utterance and Frame-Level Reference Signals

<https://doi.org/10.1109/ICASSP39728.2021.9413359>

AUTHORS: M. Ge, C. Xu, L. Wang, E. S. Chng, J. Dang and H. Li

HIGHLIGHT: We propose a speaker extraction technique, that performs in multiple stages to take full advantage of short reference speech sample.

1224, TITLE: Neural Network-Based Virtual Microphone Estimator

<https://doi.org/10.1109/ICASSP39728.2021.9415121>

AUTHORS: T. Ochiai, M. Delcroix, T. Nakatani, R. Ikeshita, K. Kinoshita and S. Araki

HIGHLIGHT: In this paper, as an alternative approach, we propose a neural network-based virtual microphone estimator (NN-VME).

1225, TITLE: Joint Maximum Likelihood Estimation of Power Spectral Densities and Relative Acoustic Transfer Functions for Acoustic Beamforming

<https://doi.org/10.1109/ICASSP39728.2021.9414252>

AUTHORS: P. Hoang, Z. -H. Tan, J. M. de Haan and J. Jensen

HIGHLIGHT: We use historical results to derive joint ML estimates (MLEs) of the RATFs and PSDs in the context of acoustic beam-forming.

1226, TITLE: Cue-Preserving MMSE Filter with Bayesian SNR Marginalization for Binaural Speech Enhancement

<https://doi.org/10.1109/ICASSP39728.2021.9414956>

AUTHORS: S. Thaleiser and G. Enzner

HIGHLIGHT: In this paper, we translate the single-channel approach into a binaural Bayesian SNR marginalization, based on a binaural a-priori SNR definition and a related hyperprior.

1227, TITLE: Blind and Neural Network-Guided Convolutional Beamformer for Joint Denoising, Dereverberation, and Source Separation

<https://doi.org/10.1109/ICASSP39728.2021.9414264>

AUTHORS: T. Nakatani, R. Ikeshita, K. Kinoshita, H. Sawada and S. Araki

HIGHLIGHT: This paper proposes an approach for optimizing a Convolutional BeamFormer (CBF) that can jointly perform denoising (DN), dereverberation (DR), and source separation (SS).

1228, TITLE: Real-Time Speech Enhancement for Mobile Communication Based on Dual-Channel Complex Spectral Mapping

<https://doi.org/10.1109/ICASSP39728.2021.9414346>

AUTHORS: K. Tan, X. Zhang and D. Wang

HIGHLIGHT: This paper proposes a novel approach to real-time speech enhancement for dual-microphone mobile phones.

1229, TITLE: Don't Shoot Butterfly with Rifles: Multi-Channel Continuous Speech Separation with Early Exit Transformer

<https://doi.org/10.1109/ICASSP39728.2021.9413933>

AUTHORS: S. Chen et al.

HIGHLIGHT: To deal with this problem, we propose an early exit mechanism, which enables the Transformer model to handle different cases with adaptive depth.

1230, TITLE: Double Multi-Head Attention for Speaker Verification

<https://doi.org/10.1109/ICASSP39728.2021.9414877>

AUTHORS: M. India, P. Safari and J. Hernando

HIGHLIGHT: In this paper we present Double Multi-Head Attention (MHA) pooling, which extends our previous approach based on Self MHA.

1231, TITLE: Graph Attention Networks for Speaker Verification

<https://doi.org/10.1109/ICASSP39728.2021.9414057>

AUTHORS: J. -w. Jung, H. -S. Heo, H. -J. Yu and J. S. Chung

HIGHLIGHT: This work presents a novel back-end framework for speaker verification using graph attention networks.

1232, TITLE: Memory Layers with Multi-Head Attention Mechanisms for Text-Dependent Speaker Verification

<https://doi.org/10.1109/ICASSP39728.2021.9414859>

AUTHORS: V. Mingote, A. Miguel, A. Ortega and E. Lleida

HIGHLIGHT: In this paper, we explore an approach based on memory layers and multi-head attention mechanisms to improve in an efficient way the performance of text-dependent speaker verification (SV) systems.

1233, TITLE: FoolHD: Fooling Speaker Identification by Highly Imperceptible Adversarial Disturbances

<https://doi.org/10.1109/ICASSP39728.2021.9413760>

AUTHORS: A. S. Shamsabadi, F. S. Teixeira, A. Abad, B. Raj, A. Cavallaro and I. Trancoso

HIGHLIGHT: In this work, we propose a white-box steganography-inspired adversarial attack that generates imperceptible adversarial perturbations against a speaker identification model.

1234, TITLE: Adversarial Defense for Deep Speaker Recognition Using Hybrid Adversarial Training

<https://doi.org/10.1109/ICASSP39728.2021.9414843>

AUTHORS: M. Pal, A. Jati, R. Peri, C. -C. Hsu, W. AbdAlmageed and S. Narayanan

HIGHLIGHT: To address this concern, in this work, we propose a new defense mechanism based on a hybrid adversarial training (HAT) setup.

1235, TITLE: DEAAAN: Disentangled Embedding and Adversarial Adaptation Network for Robust Speaker Representation Learning

<https://doi.org/10.1109/ICASSP39728.2021.9413403>

AUTHORS: M. Sang, W. Xia and J. H. L. Hansen

HIGHLIGHT: In this study, we propose a novel framework to disentangle speaker-related and domain-specific features and apply domain adaptation on the speaker-related feature space solely.

- 1236, TITLE: Playing a Part: Speaker Verification at the movies  
<https://doi.org/10.1109/ICASSP39728.2021.9413815>  
AUTHORS: A. Brown, J. Huh, A. Nagrani, J. S. Chung and A. Zisserman  
HIGHLIGHT: The goal of this work is to investigate the performance of popular speaker recognition models on speech segments from movies, where often actors intentionally disguise their voice to play a character.
- 1237, TITLE: Small Footprint Text-Independent Speaker Verification For Embedded Systems  
<https://doi.org/10.1109/ICASSP39728.2021.9413564>  
AUTHORS: J. Balian, R. Tavarone, M. Poumeyrol and A. Coucke  
HIGHLIGHT: In this work, we present a two-stage model architecture orders of magnitude smaller than common solutions (237.5K learning parameters, 11.5MFLOPS) reaching a competitive result of 3.31% Equal Error Rate (EER) on the well established VoxCeleb1 verification test set.
- 1238, TITLE: ASV-SUBTOOLS: Open Source Toolkit for Automatic Speaker Verification  
<https://doi.org/10.1109/ICASSP39728.2021.9414676>  
AUTHORS: F. Tong et al.  
HIGHLIGHT: In this paper, we introduce a new open source toolkit for automatic speaker verification (ASV), named ASV-Subtools.
- 1239, TITLE: DEEPTALK: Vocal Style Encoding for Speaker Recognition and Speech Synthesis  
<https://doi.org/10.1109/ICASSP39728.2021.9414169>  
AUTHORS: A. Chowdhury, A. Ross and P. David  
HIGHLIGHT: In this work, we propose a prosody encoding network called DeepTalk for extracting vocal style features directly from raw audio data.
- 1240, TITLE: A Multi-View Approach to Audio-Visual Speaker Verification  
<https://doi.org/10.1109/ICASSP39728.2021.9414260>  
AUTHORS: L. Sari, K. Singh, J. Zhou, L. Torresani, N. Singhal and Y. Saraf  
HIGHLIGHT: As these methods lack the ability to do cross-modal verification, we introduce a multi-view model which uses a shared classifier to map audio and video into the same space.
- 1241, TITLE: Top-Down Attention in End-to-End Spoken Language Understanding  
<https://doi.org/10.1109/ICASSP39728.2021.9414313>  
AUTHORS: Y. Chen et al.  
HIGHLIGHT: Based on this insight, we propose Top-Down SLU (TD-SLU), a new transformer-based E2E SLU model that uses top-down attention and an attention gate to fuse high-level NLU features with low-level ASR features, which leads to a better optimization of both tasks.
- 1242, TITLE: Fine-Tuning of Pre-Trained End-to-End Speech Recognition with Generative Adversarial Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9413703>  
AUTHORS: M. A. Haidar and M. Rezagholizadeh  
HIGHLIGHT: In this paper, we introduce a novel framework for fine-tuning a pre-trained ASR model using the GAN objective where the ASR model acts as a generator and a discriminator tries to distinguish the ASR output from the real data.
- 1243, TITLE: A General Multi-Task Learning Framework to Leverage Text Data for Speech to Text Tasks  
<https://doi.org/10.1109/ICASSP39728.2021.9415058>  
AUTHORS: Y. Tang, J. Pino, C. Wang, X. Ma and D. Genzel  
HIGHLIGHT: In this study, we propose a general multi-task learning framework to leverage text data for ASR and ST tasks.
- 1244, TITLE: Gaussian Kernelized Self-Attention for Long Sequence Data and its Application to CTC-Based Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9413493>  
AUTHORS: Y. Kashiwagi, E. Tsunoo and S. Watanabe  
HIGHLIGHT: To mitigate this mismatch, we propose a new architecture, which is a variant of the Gaussian kernel, which itself is a shift-invariant kernel.
- 1245, TITLE: Lattice-Free Mmi Adaptation of Self-Supervised Pretrained Acoustic Models  
<https://doi.org/10.1109/ICASSP39728.2021.9414741>  
AUTHORS: A. Vyas, S. Madikeri and H. Bourlard  
HIGHLIGHT: In this work, we propose lattice-free MMI (LFMMI) for supervised adaptation of self-supervised pretrained acoustic model.

- 1246, TITLE: Intermediate Loss Regularization for CTC-Based Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9414594>  
AUTHORS: J. Lee and S. Watanabe  
HIGHLIGHT: We present a simple and efficient auxiliary loss function for automatic speech recognition (ASR) based on the connectionist temporal classification (CTC) objective.
- 1247, TITLE: Code-Switch Speech Rescoring with Monolingual Data  
<https://doi.org/10.1109/ICASSP39728.2021.9414158>  
AUTHORS: G. Liu and L. Cao  
HIGHLIGHT: In this paper, we focus on the code-switch speech recognition in mainland China, which is obviously different from the Hong Kong and Southeast Asia area in linguistic characteristics.
- 1248, TITLE: Mixture of Informed Experts for Multilingual Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9414379>  
AUTHORS: N. Gaur et al.  
HIGHLIGHT: We introduce a novel variant of this approach, ?informed experts?, which attempts to tackle inter-task conflicts by eliminating gradients from other tasks in these task-specific parameters.
- 1249, TITLE: Reducing Spelling Inconsistencies in Code-Switching ASR Using Contextualized CTC Loss  
<https://doi.org/10.1109/ICASSP39728.2021.9413806>  
AUTHORS: B. Naowarat, T. Kongthaworn, K. Karunratanakul, S. H. Wu and E. Chuangsuwanich  
HIGHLIGHT: We propose Contextualized Connectionist Temporal Classification (CCTC) loss to encourage spelling consistencies of a character-based non-autoregressive ASR which allows for faster inference.
- 1250, TITLE: Multi-Dialect Speech Recognition in English Using Attention on Ensemble of Experts  
<https://doi.org/10.1109/ICASSP39728.2021.9413952>  
AUTHORS: A. Das, K. Kumar and J. Wu  
HIGHLIGHT: In this study, we further explore using a single model for multi-dialect speech recognition using ensemble modeling.
- 1251, TITLE: Decoupling Pronunciation and Language for End-to-End Code-Switching Automatic Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9414428>  
AUTHORS: S. Zhang, J. Yi, Z. Tian, Y. Bai, J. Tao and Z. wen  
HIGHLIGHT: In this paper, we propose a decoupled transformer model to use mono-lingual paired data and unpaired text data to alleviate the problem of code-switching data shortage.
- 1252, TITLE: AISpeech-SJTU Accent Identification System for the Accented English Speech Recognition Challenge  
<https://doi.org/10.1109/ICASSP39728.2021.9414292>  
AUTHORS: H. Huang, X. Xiang, Y. Yang, R. Ma and Y. Qian  
HIGHLIGHT: This paper describes the AISpeech-SJTU system for the accent identification track of the Interspeech-2020 Accented English Speech Recognition Challenge.
- 1253, TITLE: Meta-Learning for Low-Resource Speech Emotion Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9414373>  
AUTHORS: S. Chopra, P. Mathur, R. Sawhney and R. R. Shah  
HIGHLIGHT: Probing the learning process of generalized representations across languages, we propose a meta-learning approach for low-resource speech emotion recognition.
- 1254, TITLE: Progressive Co-Teaching for Ambiguous Speech Emotion Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9414494>  
AUTHORS: Y. Yin, Y. Gu, L. Yao, Y. Zhou, X. Liang and H. Zhang  
HIGHLIGHT: Motivated by human and animal learning studies, we propose a novel method named Progressive Co-teaching (PCT) to learn speech emotion features from simple to difficult.
- 1255, TITLE: Emotion Recognition by Fusing Time Synchronous and Time Asynchronous Representations  
<https://doi.org/10.1109/ICASSP39728.2021.9414880>  
AUTHORS: W. Wu, C. Zhang and P. C. Woodland  
HIGHLIGHT: In this paper, a novel two-branch neural network model structure is proposed for multimodal emotion recognition, which consists of a time synchronous branch (TSB) and a time asynchronous branch (TAB).



1256, TITLE: Speech Emotion Recognition Based on Listener Adaptive Models  
<https://doi.org/10.1109/ICASSP39728.2021.9414698>  
AUTHORS: A. Ando et al.  
HIGHLIGHT: In order to mitigate this problem, we propose a Listener Adaptive (LA) model that reflects emotion recognition criteria of each listener.

1257, TITLE: Speech Emotion Recognition Using Semantic Information  
<https://doi.org/10.1109/ICASSP39728.2021.9414866>  
AUTHORS: P. Tzirakis, A. Nguyen, S. Zafeiriou and B. W. Schuller  
HIGHLIGHT: In this paper, we propose a novel framework that can capture both the semantic and the paralinguistic information in the signal.

1258, TITLE: Compact Graph Architecture for Speech Emotion Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9413876>  
AUTHORS: A. Shirian and T. Guha  
HIGHLIGHT: We propose a deep graph approach to address the task of speech emotion recognition.

1259, TITLE: A Novel end-to-end Speech Emotion Recognition Network with Stacked Transformer Layers  
<https://doi.org/10.1109/ICASSP39728.2021.9414314>  
AUTHORS: X. Wang, M. Wang, W. Qi, W. Su, X. Wang and H. Zhou  
HIGHLIGHT: In contrast with those previous works, herein we propose a novel strategy for global SER feature extraction by applying an additional enhancement module on top of the current SER pipeline.

1260, TITLE: A Novel Attention-Based Gated Recurrent Unit and its Efficacy in Speech Emotion Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9414489>  
AUTHORS: S. T. Rajamani, K. T. Rajamani, A. Mallol-Ragolta, S. Liu and B. Schuller  
HIGHLIGHT: In this work, we explore using diverse activation functions within GRU and bi-directional GRU (BiGRU) cells in the context of speech emotion recognition (SER).

1261, TITLE: MAEC: Multi-Instance Learning with an Adversarial Auto-Encoder-Based Classifier for Speech Emotion Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9413640>  
AUTHORS: C. Fu, C. Liu, C. T. Ishi and H. Ishiguro  
HIGHLIGHT: In this paper, we propose an adversarial auto-encoder-based classifier, which can regularize the distribution of latent representation to smooth the boundaries among categories.

1262, TITLE: Representation Learning with Spectro-Temporal-Channel Attention for Speech Emotion Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9414006>  
AUTHORS: L. Guo, L. Wang, C. Xu, J. Dang, E. S. Chng and H. Li  
HIGHLIGHT: In this paper, we propose an attention module, named spectro-temporal-channel (STC) attention module that is integrated with CNN to improve representation learning ability.

1263, TITLE: Speech Emotion Recognition Using Quaternion Convolutional Neural Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414248>  
AUTHORS: A. Muppidi and M. Radfar  
HIGHLIGHT: Our paper addresses this problem by proposing a quaternion convolutional neural network (QCNN) based speech emotion recognition (SER) model in which Mel-spectrogram features of speech signals are encoded in an RGB quaternion domain.

1264, TITLE: Domain-Adversarial Autoencoder with Attention Based Feature Level Fusion for Speech Emotion Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9413710>  
AUTHORS: Y. Gao, J. Liu, L. Wang and J. Dang  
HIGHLIGHT: In this paper, we propose a domain-adversarial autoencoder to extract discriminative representations for SER.

1265, TITLE: Speech Emotion Recognition with Multiscale Area Attention and Data Augmentation  
<https://doi.org/10.1109/ICASSP39728.2021.9414635>  
AUTHORS: M. Xu, F. Zhang, X. Cui and W. Zhang  
HIGHLIGHT: In this paper, we apply multiscale area attention in a deep convolutional neural network to attend emotional characteristics with varied granularities and therefore the classifier can benefit from an ensemble of attentions with different scales.

- 1266, TITLE: CopyPaste: An Augmentation Method for Speech Emotion Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9415077>  
AUTHORS: R. Pappagari, J. Villalba, P. Zelasko, L. Moro-Velazquez and N. Dehak  
HIGHLIGHT: This study proposes CopyPaste, a perceptually motivated novel augmentation procedure for SER.
- 1267, TITLE: Contrastive Unsupervised Learning for Speech Emotion Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9413910>  
AUTHORS: M. Li et al.  
HIGHLIGHT: To circumvent this problem, we investigate how unsupervised representation learning on unlabeled datasets can benefit SER.
- 1268, TITLE: Hierarchical Network Based on the Fusion of Static and Dynamic Features for Speech Emotion Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9414540>  
AUTHORS: Q. Cao, M. Hou, B. Chen, Z. Zhang and G. Lu  
HIGHLIGHT: In this paper, we propose a novel hierarchical network called HNSD that can efficiently integrate the static and dynamic features for SER.
- 1269, TITLE: Multimodal Emotion Recognition with Capsule Graph Convolutional Based Representation Fusion  
<https://doi.org/10.1109/ICASSP39728.2021.9413608>  
AUTHORS: J. Liu et al.  
HIGHLIGHT: In this paper, we propose a novel representation fusion method, Capsule Graph Convolutional Network (CapsGCN).
- 1270, TITLE: Disentanglement for Audio-Visual Emotion Recognition Using Multitask Setup  
<https://doi.org/10.1109/ICASSP39728.2021.9414705>  
AUTHORS: R. Peri, S. Parthasarathy, C. Bradshaw and S. Sundaram  
HIGHLIGHT: In particular, we developed a multitask framework to extract low-dimensional embeddings that aim to capture emotion specific information, while containing minimal information related to person identity.
- 1271, TITLE: Data Augmentation with Signal Companding for Detection of Logical Access Attacks  
<https://doi.org/10.1109/ICASSP39728.2021.9413501>  
AUTHORS: R. K. Das, J. Yang and H. Li  
HIGHLIGHT: In this work, we propose a novel data augmentation technique using a-law and mu-law based signal companding.
- 1272, TITLE: Replay and Synthetic Speech Detection with Res2Net Architecture  
<https://doi.org/10.1109/ICASSP39728.2021.9413828>  
AUTHORS: X. Li et al.  
HIGHLIGHT: This work proposes to leverage a novel model structure, so-called Res2Net, to improve the anti-spoofing countermeasure's generalizability.
- 1273, TITLE: A Capsule Network Based Approach for Detection of Audio Spoofing Attacks  
<https://doi.org/10.1109/ICASSP39728.2021.9414670>  
AUTHORS: A. Luo, E. Li, Y. Liu, X. Kang and Z. J. Wang  
HIGHLIGHT: In this paper, as the first attempt, we introduce a capsule network to enhance the generalization of the detection system.
- 1274, TITLE: Cross-Teager Energy Cepstral Coefficients for Replay Spoof Detection on Voice Assistants  
<https://doi.org/10.1109/ICASSP39728.2021.9414847>  
AUTHORS: R. Acharya, H. Kotta, A. T. Patil and H. A. Patil  
HIGHLIGHT: The key idea of this work is optimal channel selection based on maximum cross-energies from a multichannel input, which is suitable for SSD task.
- 1275, TITLE: End-to-End anti-spoofing with RawNet2  
<https://doi.org/10.1109/ICASSP39728.2021.9414234>  
AUTHORS: H. Tak, J. Patino, M. Todisco, A. Nautsch, N. Evans and A. Larcher  
HIGHLIGHT: We describe modifications made to the original RawNet2 architecture so that it can be applied to anti-spoofing.

- 1276, TITLE:       Replay-Attack Detection Using Features With Adaptive Spectro-Temporal Resolution  
<https://doi.org/10.1109/ICASSP39728.2021.9414250>  
AUTHORS:       M. Liu, L. Wang, K. A. Lee, X. Chen and J. Dang  
HIGHLIGHT:       In this paper, an adaptive spectro-temporal resolution is proposed to obtain the optimal scale in the feature space: the frequency resolution is adaptive to frequency discrimination, while the temporal resolution is adaptive to continuous phones.
- 1277, TITLE:       Improving Identification of System-Directed Speech Utterances by Deep Learning of ASR-Based Word Embeddings and Confidence Metrics  
<https://doi.org/10.1109/ICASSP39728.2021.9414330>  
AUTHORS:       V. Vilaysouk, A. Nour-Eldin and D. Connolly  
HIGHLIGHT:       In this paper, we extend our previous work on the detection of system-directed speech utterances.
- 1278, TITLE:       BLSTM-Based Confidence Estimation for End-to-End Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9414977>  
AUTHORS:       A. Ogawa, N. Tawara, T. Kano and M. Delcroix  
HIGHLIGHT:       In this study, we perform confidence estimation for end-to-end (E2E) ASR hypotheses.
- 1279, TITLE:       Confidence Estimation for Attention-Based Sequence-to-Sequence Models for Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9414920>  
AUTHORS:       Q. Li et al.  
HIGHLIGHT:       In this paper, we first examine how some commonly used regularisation methods influence the softmax-based confidence scores and study the overconfident behaviour of end-to-end models. Then we propose a lightweight and effective approach named confidence estimation module (CEM) on top of an existing end-to-end ASR model.
- 1280, TITLE:       Learning Word-Level Confidence for Subword End-To-End ASR  
<https://doi.org/10.1109/ICASSP39728.2021.9413966>  
AUTHORS:       D. Qiu et al.  
HIGHLIGHT:       This paper proposes and studies two confidence models of increasing complexity to solve this problem.
- 1281, TITLE:       Neural Utterance Confidence Measure for RNN-Transducers and Two Pass Models  
<https://doi.org/10.1109/ICASSP39728.2021.9414467>  
AUTHORS:       A. Gupta et al.  
HIGHLIGHT:       In this paper, we propose methods to compute confidence score on the predictions made by an end-to-end speech recognition model in a 2-pass framework.
- 1282, TITLE:       Detecting Adversarial Attacks on Audiovisual Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9413661>  
AUTHORS:       P. Ma, S. Petridis and M. Pantic  
HIGHLIGHT:       In this work, we propose an efficient and straightforward detection method based on the temporal correlation between audio and video streams.
- 1283, TITLE:       REDAT: Accent-Invariant Representation for End-To-End ASR by Domain Adversarial Training with Relabeling  
<https://doi.org/10.1109/ICASSP39728.2021.9414291>  
AUTHORS:       H. Hu et al.  
HIGHLIGHT:       Motivated by the proof of equivalence, we introduce reDAT, a novel technique based on DAT, which relabels data using either unsupervised clustering or soft labels.
- 1284, TITLE:       AISpeech-SJTU ASR System for the Accented English Speech Recognition Challenge  
<https://doi.org/10.1109/ICASSP39728.2021.9414471>  
AUTHORS:       T. Tan, Y. Lu, R. Ma, S. Zhu, J. Guo and Y. Qian  
HIGHLIGHT:       This paper describes the AISpeech-SJTU ASR system for the Interspeech-2020 Accented English Speech Recognition Challenge (AESRC).
- 1285, TITLE:       End-To-End Multi-Accent Speech Recognition with Unsupervised Accent Modelling  
<https://doi.org/10.1109/ICASSP39728.2021.9414833>  
AUTHORS:       S. Li, B. Ouyang, D. Liao, S. Xia, L. Li and Q. Hong  
HIGHLIGHT:       In order to grapple with such an issue, we first investigate and improve the current mainstream end-to-end multi-accent speech recognition technologies. In addition, we propose two unsupervised accent modelling methods, which convert

accent information into a global embedding, and use it to improve the performance of the end-to-end multi-accent speech recognition systems.

1286, TITLE: A Comparative Study of Acoustic and Linguistic Features Classification for Alzheimer's Disease Detection  
<https://doi.org/10.1109/ICASSP39728.2021.9414147>  
AUTHORS: J. Li et al.  
HIGHLIGHT: This paper presents a comparative study of different acoustic and linguistic features for the AD detection using various classifiers.

1287, TITLE: Synthesis of New Words for Improved Dysarthric Speech Recognition on an Expanded Vocabulary  
<https://doi.org/10.1109/ICASSP39728.2021.9414869>  
AUTHORS: J. Harvill, D. Issa, M. Hasegawa-Johnson and C. Yoo  
HIGHLIGHT: In this paper, we propose a data augmentation method using voice conversion that allows dysarthric ASR systems to accurately recognize words outside of the training set vocabulary.

1288, TITLE: Development of the Cuhk Elderly Speech Recognition System for Neurocognitive Disorder Detection Using the Dementiabank Corpus  
<https://doi.org/10.1109/ICASSP39728.2021.9413634>  
AUTHORS: Z. Ye et al.  
HIGHLIGHT: This paper presents the development of a state-of-the-art automatic speech recognition (ASR) system built on the Dementia-Bank Pitt corpus for automatic NCD detection.

1289, TITLE: Portable Photoglottography for Monitoring Vocal Fold Vibrations in Speech Production  
<https://doi.org/10.1109/ICASSP39728.2021.9413770>  
AUTHORS: Y. Chi, K. Honda and J. Wei  
HIGHLIGHT: This paper is to realize a portable PGG (P-PGG) module with an audio interface to record glottal and speech waveforms simultaneously with ease.

1290, TITLE: Improving Ultrasound Tongue Contour Extraction Using U-Net and Shape Consistency-Based Regularizer  
<https://doi.org/10.1109/ICASSP39728.2021.9414420>  
AUTHORS: M. Feng, Y. Wang, K. Xu, H. Wang and B. Ding  
HIGHLIGHT: To address the faint or missing contours in the sequence, we explore the shape consistency-based regularizer, which can take sequential information into account.

1291, TITLE: Impact of Speaking Rate on the Source Filter Interaction in Speech: A Study  
<https://doi.org/10.1109/ICASSP39728.2021.9414221>  
AUTHORS: T. Purohit, A. R. MV and P. Kumar Ghosh  
HIGHLIGHT: In this work, we examine how the drop in pitch alters when such a VCV sequence is spoken at three different speaking rates - slow, normal and fast.

1292, TITLE: A Two-Stage Deep Modeling Approach to Articulatory Inversion  
<https://doi.org/10.1109/ICASSP39728.2021.9413742>  
AUTHORS: A. S. Shahrehabaki, N. Olfati, A. S. Imran, M. Hallstein Johnsen, S. M. Siniscalchi and T. Svendsen  
HIGHLIGHT: This paper proposes a two-stage deep feed-forward neural network (DNN) to tackle the acoustic-to-articulatory inversion (AAI) problem.

1293, TITLE: Acoustic-to-Articulatory Inversion for Dysarthric Speech by Using Cross-Corpus Acoustic-Articulatory Data  
<https://doi.org/10.1109/ICASSP39728.2021.9413625>  
AUTHORS: S. K. Maharana et al.  
HIGHLIGHT: In this work, we focus on estimating articulatory movements from acoustic features, known as acoustic-to-articulatory inversion (AAI), for dysarthric patients with amyotrophic lateral sclerosis (ALS).

1294, TITLE: Speaking Rate and Tonal Realization in Mandarin Chinese: What Can We Learn From Large Speech Corpora?  
<https://doi.org/10.1109/ICASSP39728.2021.9413838>  
AUTHORS: J. Yuan and K. Church  
HIGHLIGHT: Our analysis found two differences for slower speaking rates: (1) lower "static" tones and (2) more change for "dynamic" tones.

1295, TITLE: Humanacgan: Conditional Generative Adversarial Network with Human-Based Auxiliary Classifier and its Evaluation in Phoneme Perception

<https://doi.org/10.1109/ICASSP39728.2021.9413363>

AUTHORS: Y. Ueda, K. Fujii, Y. Saito, S. Takamichi, Y. Baba and H. Saruwatari

HIGHLIGHT: We propose a conditional generative adversarial network (GAN) incorporating humans' perceptual evaluations.

1296, TITLE: Improving Audio Anomalies Recognition Using Temporal Convolutional Attention Networks

<https://doi.org/10.1109/ICASSP39728.2021.9414611>

AUTHORS: Q. Huang and T. Hain

HIGHLIGHT: In this paper, a novel approach using a temporal convolutional attention network (TCAN) is proposed to tackle this problem.

1297, TITLE: Generative Speech Coding with Predictive Variance Regularization

<https://doi.org/10.1109/ICASSP39728.2021.9415120>

AUTHORS: W. B. Kleijn et al.

HIGHLIGHT: We introduce predictive-variance regularization to reduce the sensitivity to outliers, resulting in a significant increase in performance.

1298, TITLE: How to Make Text-to-Speech System Pronounce "Voldemort": an Experimental Approach of Foreign Word Phonemization in Vietnamese

<https://doi.org/10.1109/ICASSP39728.2021.9414386>

AUTHORS: D. -K. MAC, V. -H. NGUYEN, D. -N. NGUYEN and K. -A. NGUYEN

HIGHLIGHT: Generating foreign words is one of the hardest tasks for any speech synthesis systems. This work deal with this problem in the case of Vietnamese, a low-resourced language, following an experimental approach.

1299, TITLE: How Similar or Different is Rakugo Speech Synthesizer to Professional Performers?

<https://doi.org/10.1109/ICASSP39728.2021.9414175>

AUTHORS: S. Kato, Y. Yasuda, X. Wang, E. Cooper and J. Yamagishi

HIGHLIGHT: In this paper, we propose a novel evaluation methodology using synthesized rakugo speech and real rakugo speech uttered by professional performers of three different ranks.

1300, TITLE: Dnsmos: A Non-Intrusive Perceptual Objective Speech Quality Metric to Evaluate Noise Suppressors

<https://doi.org/10.1109/ICASSP39728.2021.9414878>

AUTHORS: C. K. A. Reddy, V. Gopal and R. Cutler

HIGHLIGHT: This paper introduces a multi-stage self-teaching based perceptual objective metric that is designed to evaluate noise suppressors.

1301, TITLE: A Causal Deep Learning Framework for Classifying Phonemes in Cochlear Implants

<https://doi.org/10.1109/ICASSP39728.2021.9413986>

AUTHORS: K. Chu, L. Collins and B. Mainsah

HIGHLIGHT: In this paper, we propose a causal deep learning framework for classifying phonemes using features extracted at the time-frequency resolution of a CI processor.

1302, TITLE: Minimum Bayes Risk Training for End-to-End Speaker-Attributed ASR

<https://doi.org/10.1109/ICASSP39728.2021.9415062>

AUTHORS: N. Kanda et al.

HIGHLIGHT: In this paper, we propose a speaker-attributed minimum Bayes risk (SA-MBR) training method where the parameters are trained to directly minimize the expected SA-WER over the training data.

1303, TITLE: Mutually-Constrained Monotonic Multihead Attention for Online ASR

<https://doi.org/10.1109/ICASSP39728.2021.9413862>

AUTHORS: J. Song, H. Shim and E. Yang

HIGHLIGHT: In this paper, we remove the discrepancy between training and test phases by considering, in the training of MMA, the interactions across multiple heads that will occur in the test time.

1304, TITLE: The use of Voice Source Features for Sung Speech Recognition

<https://doi.org/10.1109/ICASSP39728.2021.9414950>

AUTHORS: G. R. Dabike and J. Barker

HIGHLIGHT: In this paper, we ask whether vocal source features (pitch, shimmer, jitter, etc) can improve the performance of automatic sung speech recognition, arguing that conclusions previously drawn from spoken speech studies may not be valid in the sung speech domain.

- 1305, TITLE: A Parallelizable Lattice Rescoring Strategy with Neural Language Models  
<https://doi.org/10.1109/ICASSP39728.2021.9414714>  
AUTHORS: K. Li, D. Povey and S. Khudanpur  
HIGHLIGHT: This paper proposes a parallel computation strategy and a posterior-based lattice expansion algorithm for efficient lattice rescoring with neural language models (LMs) for automatic speech recognition.
- 1306, TITLE: Decentralizing Feature Extraction with Quantum Convolutional Neural Network for Automatic Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9413453>  
AUTHORS: C. -H. H. Yang et al.  
HIGHLIGHT: We propose a novel decentralized feature extraction approach in federated learning to address privacy-preservation issues for speech recognition.
- 1307, TITLE: Cif-Based Collaborative Decoding for End-to-End Contextual Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9415054>  
AUTHORS: M. Han, L. Dong, S. Zhou and B. Xu  
HIGHLIGHT: In this paper, we focus on incorporating contextual information into the continuous integrate-and-fire (CIF) based model that supports contextual biasing in a more controllable fashion.
- 1308, TITLE: Hubert: How Much Can a Bad Teacher Benefit ASR Pre-Training?  
<https://doi.org/10.1109/ICASSP39728.2021.9414460>  
AUTHORS: W. -N. Hsu, Y. -H. H. Tsai, B. Bolte, R. Salakhutdinov and A. Mohamed  
HIGHLIGHT: In this paper, we propose the Hidden-Unit BERT (HUBERT) model which utilizes a cheap k-means clustering step to provide aligned target labels for pre-training of a BERT model.
- 1309, TITLE: A Further Study of Unsupervised Pretraining for Transformer Based Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9414539>  
AUTHORS: D. Jiang et al.  
HIGHLIGHT: In this paper, we conduct a further study on MPC and focus on three important aspects: the effect of pretraining data speaking style, its extension on streaming model, and strategies for better transferring learned knowledge from pretraining stage to downstream tasks.
- 1310, TITLE: Pre-Training Transformer Decoder for End-to-End ASR Model with Unpaired Text Data  
<https://doi.org/10.1109/ICASSP39728.2021.9414080>  
AUTHORS: C. Gao, G. Cheng, R. Yang, H. Zhu, P. Zhang and Y. Yan  
HIGHLIGHT: This paper presents a method to pre-train transformer-based encoder-decoder automatic speech recognition (ASR) models using sufficient target-domain text.
- 1311, TITLE: Semi-Supervised Speech Recognition Via Graph-Based Temporal Classification  
<https://doi.org/10.1109/ICASSP39728.2021.9414058>  
AUTHORS: N. Moritz, T. Hori and J. L. Roux  
HIGHLIGHT: In this paper, we propose a generalized form of the connectionist temporal classification (CTC) objective that accepts a graph representation of the training labels.
- 1312, TITLE: Unsupervised Domain Adaptation for Speech Recognition via Uncertainty Driven Self-Training  
<https://doi.org/10.1109/ICASSP39728.2021.9414299>  
AUTHORS: S. Khurana, N. Moritz, T. Hori and J. L. Roux  
HIGHLIGHT: In this paper, we show that self-training (ST) combined with an uncertainty-based pseudo-label filtering approach can be effectively used for domain adaptation.
- 1313, TITLE: Improving Streaming Automatic Speech Recognition with Non-Streaming Model Distillation on Unsupervised Data  
<https://doi.org/10.1109/ICASSP39728.2021.9413692>  
AUTHORS: T. Doure et al.  
HIGHLIGHT: We propose a novel and effective learning method by leveraging a non-streaming ASR model as a teacher to generate transcripts on an arbitrarily large data set, which is then used to distill knowledge into streaming ASR models.
- 1314, TITLE: Speech Bert Embedding for Improving Prosody in Neural TTS  
<https://doi.org/10.1109/ICASSP39728.2021.9413864>  
AUTHORS: L. Chen, Y. Deng, X. Wang, F. K. Soong and L. He

**HIGHLIGHT:** This paper presents a speech BERT model to extract embedded prosody information in speech segments for improving the prosody of synthesized speech in neural text-to-speech (TTS).

1315, **TITLE:** Bi-Level Style and Prosody Decoupling Modeling for Personalized End-to-End Speech Synthesis

<https://doi.org/10.1109/ICASSP39728.2021.9414422>

**AUTHORS:** R. Fu, J. Tao, Z. Wen, J. Yi, T. Wang and C. Qiang

**HIGHLIGHT:** In this paper, we present a bi-level function decoupling framework to realise separate modeling and controlling for solving above problems.

1316, **TITLE:** Prosodic Representation Learning and Contextual Sampling for Neural Text-to-Speech

<https://doi.org/10.1109/ICASSP39728.2021.9413696>

**AUTHORS:** S. Karlapati et al.

**HIGHLIGHT:** In this paper, we introduce Kathaka, a model trained with a novel two-stage training process for neural speech synthesis with contextually appropriate prosody.

1317, **TITLE:** Camp: A Two-Stage Approach to Modelling Prosody in Context

<https://doi.org/10.1109/ICASSP39728.2021.9414413>

**AUTHORS:** Z. Hodari et al.

**HIGHLIGHT:** There are two major issues faced when modelling prosody: (1) prosody varies at a slower rate compared with other content in the acoustic signal (e.g. segmental information and background noise); (2) determining appropriate prosody without sufficient context is an ill-posed problem. In this paper, we propose solutions to both these issues.

1318, **TITLE:** Unsupervised Learning for Multi-Style Speech Synthesis with Limited Data

<https://doi.org/10.1109/ICASSP39728.2021.9414220>

**AUTHORS:** S. Liang, C. Miao, M. Chen, J. Ma, S. Wang and J. Xiao

**HIGHLIGHT:** In this paper, we present an unsupervised multi-style speech synthesis method that can be trained with limited data.

1319, **TITLE:** Fastpitch: Parallel Text-to-Speech with Pitch Prediction

<https://doi.org/10.1109/ICASSP39728.2021.9413889>

**AUTHORS:** A. Lancucki

**HIGHLIGHT:** We present FastPitch, a fully-parallel text-to-speech model based on FastSpeech, conditioned on fundamental frequency contours.

1320, **TITLE:** Low-Resource Expressive Text-To-Speech Using Data Augmentation

<https://doi.org/10.1109/ICASSP39728.2021.9413466>

**AUTHORS:** G. Huybrechts, T. Merritt, G. Comini, B. Perz, R. Shah and J. Lorenzo-Trueba

**HIGHLIGHT:** In this work, we present a novel 3-step methodology to circumvent the costly operation of recording large amounts of target data in order to build expressive style voices with as little as 15 minutes of such recordings.

1321, **TITLE:** TTS-by-TTS: TTS-Driven Data Augmentation for Fast and High-Quality Speech Synthesis

<https://doi.org/10.1109/ICASSP39728.2021.9414408>

**AUTHORS:** M. -J. Hwang, R. Yamamoto, E. Song and J. -M. Kim

**HIGHLIGHT:** In this paper, we propose a text-to-speech (TTS)-driven data augmentation method for improving the quality of a non-autoregressive (AR) TTS system.

1322, **TITLE:** A Neural Text-to-Speech Model Utilizing Broadcast Data Mixed with Background Music

<https://doi.org/10.1109/ICASSP39728.2021.9415061>

**AUTHORS:** H. Bae, J. -S. Bae, Y. -S. Joo, Y. -I. Kim and H. -Y. Cho

**HIGHLIGHT:** Therefore, we propose the following method to successfully train an end-to-end TTS model with limited broadcast data.

1323, **TITLE:** Disentangled Speaker and Language Representations Using Mutual Information Minimization and Domain Adaptation for Cross-Lingual TTS

<https://doi.org/10.1109/ICASSP39728.2021.9414226>

**AUTHORS:** D. Xin, T. Komatsu, S. Takamichi and H. Saruwatari

**HIGHLIGHT:** We propose a method for obtaining disentangled speaker and language representations via mutual information minimization and domain adaptation for cross-lingual text-to-speech (TTS) synthesis.

1324, **TITLE:** Adaspeech 2: Adaptive Text to Speech with Untranscribed Data



<https://doi.org/10.1109/ICASSP39728.2021.9414872>

AUTHORS: Y. Yan et al.

HIGHLIGHT: In this paper, we develop AdaSpeech 2, an adaptive TTS system that only leverages untranscribed speech data for adaptation.

1325, TITLE: Investigation of Fast and Efficient Methods for Multi-Speaker Modeling and Speaker Adaptation

<https://doi.org/10.1109/ICASSP39728.2021.9413396>

AUTHORS: Y. Zheng, X. Li and L. Lu

HIGHLIGHT: In this paper, we propose a novel method for fast and efficient few-shot TTS task, which is able to disentangle linguistic and speaker representations.

1326, TITLE: ICASSP 2021 Deep Noise Suppression Challenge

<https://doi.org/10.1109/ICASSP39728.2021.9415105>

AUTHORS: C. K. A. Reddy et al.

HIGHLIGHT: The Deep Noise Suppression (DNS) challenge is designed to foster innovation in the area of noise suppression to achieve superior perceptual speech quality.

1327, TITLE: ICASSP 2021 Deep Noise Suppression Challenge: Decoupling Magnitude and Phase Optimization with a Two-Stage Deep Network

<https://doi.org/10.1109/ICASSP39728.2021.9414062>

AUTHORS: A. Li, W. Liu, X. Luo, C. Zheng and X. Li

HIGHLIGHT: To this end, we propose a novel system for denoising in the complicated applications, which is mainly comprised of two pipelines, namely a two-stage network and a post-processing module.

1328, TITLE: Fullsubnet: A Full-Band and Sub-Band Fusion Model for Real-Time Single-Channel Speech Enhancement

<https://doi.org/10.1109/ICASSP39728.2021.9414177>

AUTHORS: X. Hao, X. Su, R. Horaud and X. Li

HIGHLIGHT: This paper proposes a full-band and sub-band fusion model, named as FullSubNet, for single-channel real-time speech enhancement.

1329, TITLE: Densely Connected Multi-Stage Model with Channel Wise Subband Feature for Real-Time Speech Enhancement

<https://doi.org/10.1109/ICASSP39728.2021.9413967>

AUTHORS: J. Li et al.

HIGHLIGHT: In this paper, we propose a computationally efficient real-time speech enhancement network with densely connected multi-stage structures, which progressively enhances the channel-wise subband speech.

1330, TITLE: A Modulation-Domain Loss for Neural-Network-Based Real-Time Speech Enhancement

<https://doi.org/10.1109/ICASSP39728.2021.9414965>

AUTHORS: T. Vuong, Y. Xia and R. M. Stern

HIGHLIGHT: We describe a modulation-domain loss function for deep-learning-based speech enhancement systems.

1331, TITLE: Monaural Speech Enhancement with Complex Convolutional Block Attention Module and Joint Time Frequency Losses

<https://doi.org/10.1109/ICASSP39728.2021.9414569>

AUTHORS: S. Zhao, T. H. Nguyen and B. Ma

HIGHLIGHT: In this paper, we propose a complex convolutional block attention module (CCBAM) to boost the representation power of the complex-valued convolutional layers by constructing more informative features.

1332, TITLE: Audio-Visual Speech Inpainting with Deep Learning

<https://doi.org/10.1109/ICASSP39728.2021.9413488>

AUTHORS: G. Morrone, D. Michelsanti, Z. -H. Tan and J. Jensen

HIGHLIGHT: In this paper, we present a deep-learning-based framework for audio-visual speech inpainting, i.e., the task of restoring the missing parts of an acoustic speech signal from reliable audio context and uncorrupted visual information.

1333, TITLE: Vset: A Multimodal Transformer for Visual Speech Enhancement

<https://doi.org/10.1109/ICASSP39728.2021.9414053>

AUTHORS: K. Ramesh, C. Xing, W. Wang, D. Wang and X. Chen

HIGHLIGHT: In this paper, we challenge this common belief and show that an audio-visual transformer can significantly improve AVSE performance, by learning the long-term dependency of both intra-modality and inter-modality.

- 1334, TITLE: Switching Variational Auto-Encoders for Noise-Agnostic Audio-Visual Speech Enhancement  
<https://doi.org/10.1109/ICASSP39728.2021.9414097>  
AUTHORS: M. Sadeghi and X. Alameda-Pineda  
HIGHLIGHT: In this paper, we propose to find the optimal combination of these two architectures through time.
- 1335, TITLE: Audio-Visual Speech Enhancement Method Conditioned in the Lip Motion and Speaker-Discriminative Embeddings  
<https://doi.org/10.1109/ICASSP39728.2021.9414133>  
AUTHORS: K. Ito, M. Yamamoto and K. Nagamatsu  
HIGHLIGHT: We propose an audio-visual speech enhancement (AVSE) method conditioned both on the speaker's lip motion and on speaker-discriminative embeddings.
- 1336, TITLE: Audio-Visual Speech Separation Using Cross-Modal Correspondence Loss  
<https://doi.org/10.1109/ICASSP39728.2021.9413491>  
AUTHORS: N. Makishima, M. Ihori, A. Takashima, T. Tanaka, S. Orihashi and R. Masumura  
HIGHLIGHT: We present an audio-visual speech separation learning method that considers the correspondence between the separated signals and the visual signals to reflect the speech characteristics during training.
- 1337, TITLE: Muse: Multi-Modal Target Speaker Extraction with Visual Cues  
<https://doi.org/10.1109/ICASSP39728.2021.9414023>  
AUTHORS: Z. Pan, R. Tao, C. Xu and H. Li  
HIGHLIGHT: We propose a multi-modal speaker extraction network, named MuSE, that is conditioned only on a lip image sequence.
- 1338, TITLE: An Effective Deep Embedding Learning Method Based on Dense-Residual Networks for Speaker Verification  
<https://doi.org/10.1109/ICASSP39728.2021.9413421>  
AUTHORS: Y. Liu, Y. Song, I. McLoughlin, L. Liu and L. -r. Dai  
HIGHLIGHT: In this paper, we present an effective end-to-end deep embedding learning method based on Dense-Residual networks, which combine the advantages of a densely connected convolutional network (DenseNet) and a residual network (ResNet), for speaker verification (SV).
- 1339, TITLE: Time-Domain Speaker Verification Using Temporal Convolutional Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414765>  
AUTHORS: S. Han, J. Byun and J. W. Shin  
HIGHLIGHT: In this paper, we propose a speaker verification system that takes the time-domain raw waveforms as inputs, which adopts a learnable encoder and temporal convolutional networks (TCNs) that have shown impressive performance in speech separation.
- 1340, TITLE: Towards Robust Speaker Verification with Target Speaker Enhancement  
<https://doi.org/10.1109/ICASSP39728.2021.9414017>  
AUTHORS: C. Zhang, M. Yu, C. Weng and D. Yu  
HIGHLIGHT: This paper proposes the target speaker enhancement based speaker verification network (TASE-SVNet), an all neural model that couples target speaker enhancement and speaker embedding extraction for robust speaker verification (SV).
- 1341, TITLE: A Joint Training Framework of Multi-Look Separator and Speaker Embedding Extractor for Overlapped Speech  
<https://doi.org/10.1109/ICASSP39728.2021.9414556>  
AUTHORS: N. Zheng et al.  
HIGHLIGHT: In this paper, a joint training framework of the front-end multi-look speech separator and the back-end speaker embedding extractor is proposed for multi-channel overlapped speech.
- 1342, TITLE: Cam: Context-Aware Masking for Robust Speaker Verification  
<https://doi.org/10.1109/ICASSP39728.2021.9414704>  
AUTHORS: Y. -Q. Yu, S. Zheng, H. Suo, Y. Lei and W. -J. Li  
HIGHLIGHT: In this paper, we propose context-aware masking (CAM), a novel method to extract robust speaker embedding.
- 1343, TITLE: Short-Time Spectral Aggregation for Speaker Embedding  
<https://doi.org/10.1109/ICASSP39728.2021.9414094>  
AUTHORS: Y. Tu and M. -W. Mak  
HIGHLIGHT: This paper introduces short-time spectral pooling (STSP) for better aggregation of frame-level information.

- 1344, TITLE: Contrastive Self-Supervised Learning for Text-Independent Speaker Verification  
<https://doi.org/10.1109/ICASSP39728.2021.9413351>  
AUTHORS: H. Zhang, Y. Zou and H. Wang  
HIGHLIGHT: We proposed channel-invariant loss to prevent the network from encoding the undesired channel information into the speaker representation.
- 1345, TITLE: Adversarial Defense for Automatic Speaker Verification by Cascaded Self-Supervised Learning Models  
<https://doi.org/10.1109/ICASSP39728.2021.9413737>  
AUTHORS: H. Wu, X. Li, A. T. Liu, Z. Wu, H. Meng and H. -y. Lee  
HIGHLIGHT: Hence, with the goal of effective defense in ASV against adversarial attacks, we propose a standard and attack-agnostic method based on cascaded self-supervised learning models to purify the adversarial perturbations.
- 1346, TITLE: Self-Supervised Text-Independent Speaker Verification Using Prototypical Momentum Contrastive Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9414973>  
AUTHORS: W. Xia, C. Zhang, C. Weng, M. Yu and D. Yu  
HIGHLIGHT: In this study, we investigate self-supervised representation learning for speaker verification (SV).
- 1347, TITLE: An Iterative Framework for Self-Supervised Deep Speaker Representation Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9414713>  
AUTHORS: D. Cai, W. Wang and M. Li  
HIGHLIGHT: In this paper, we propose an iterative framework for self-supervised speaker representation learning based on a deep neural network (DNN).
- 1348, TITLE: Improving Reconstruction Loss Based Speaker Embedding in Unsupervised and Semi-Supervised Scenarios  
<https://doi.org/10.1109/ICASSP39728.2021.9413739>  
AUTHORS: J. Cho, P. Zelasko, J. Villalba and N. Dehak  
HIGHLIGHT: Thus, in this paper, we evaluate speaker embeddings learned by training the spectrogram prediction network under unsupervised and semi-supervised scenarios.
- 1349, TITLE: Speech Acoustic Modelling from Raw Phase Spectrum  
<https://doi.org/10.1109/ICASSP39728.2021.9413727>  
AUTHORS: E. Loweimi, Z. Cvetkovic, P. Bell and S. Renals  
HIGHLIGHT: In this paper, we investigate the possibility and efficacy of acoustic modelling using the raw short-time phase spectrum.
- 1350, TITLE: An Investigation of Using Hybrid Modeling Units for Improving End-to-End Speech Recognition System  
<https://doi.org/10.1109/ICASSP39728.2021.9414598>  
AUTHORS: S. Chen, X. Hu, S. Li and X. Xu  
HIGHLIGHT: This paper uses a hybrid of the syllable, Chinese character, and subword as the modeling units for the end-to-end speech recognition system based on the CTC/attention multi-task learning.
- 1351, TITLE: Federated Acoustic Modeling for Automatic Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9414305>  
AUTHORS: X. Cui, S. Lu and B. Kingsbury  
HIGHLIGHT: In this paper, we investigate federated acoustic modeling using data from multiple clients.
- 1352, TITLE: Eat: Enhanced ASR-TTS for Self-Supervised Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9413375>  
AUTHORS: M. K. Baskar, L. Burget, S. Watanabe, R. F. Astudillo and J. ". Cernock?  
HIGHLIGHT: Here we propose an enhanced ASR-TTS (EAT) model that incorporates two main features: 1) The ASR?TTS direction is equipped with a language model reward to penalize the ASR hypotheses before forwarding it to TTS.
- 1353, TITLE: Neural Architecture Search for LF-MMI Trained Time Delay Neural Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9413630>  
AUTHORS: S. Hu et al.  
HIGHLIGHT: In this paper, a range of neural architecture search (NAS) techniques are used to automatically learn two types of hyper-parameters of state-of-the-art factored time delay neural networks (TDNNs): i) the left and right splicing context offsets; and ii) the dimensionality of the bottleneck linear projection at each hidden layer.

- 1354, TITLE: Hypothesis Stitcher for End-to-End Speaker-Attributed ASR on Long-Form Multi-Talker Recordings  
<https://doi.org/10.1109/ICASSP39728.2021.9414432>  
AUTHORS: X. Chang, N. Kanda, Y. Gaur, X. Wang, Z. Meng and T. Yoshioka  
HIGHLIGHT: In this work, we first apply a known decoding technique that was developed to perform single-speaker ASR for long-form audio to our E2E SA-ASR task. Then, we propose a novel method using a sequence-to-sequence model, called hypothesis stitcher.
- 1355, TITLE: Ensemble Combination between Different Time Segmentations  
<https://doi.org/10.1109/ICASSP39728.2021.9413819>  
AUTHORS: J. H. M. Wong et al.  
HIGHLIGHT: This paper proposes to generalise hypothesis-level combination, allowing the use of different audio segmentation times between the models, by splitting and re-joining the hypothesised N-best lists in time.
- 1356, TITLE: Streaming End-to-End Speech Recognition with Jointly Trained Neural Feature Enhancement  
<https://doi.org/10.1109/ICASSP39728.2021.9414117>  
AUTHORS: C. Kim, A. Garg, D. Gowda, S. Mun and C. Han  
HIGHLIGHT: In this paper, we present a streaming end-to-end speech recognition model based on Monotonic Chunkwise Attention (MoCha) jointly trained with enhancement layers.
- 1357, TITLE: Transformer in Action: A Comparative Study of Transformer-Based Acoustic Models for Large Scale Speech Recognition Applications  
<https://doi.org/10.1109/ICASSP39728.2021.9414087>  
AUTHORS: Y. Wang et al.  
HIGHLIGHT: In this paper, we summarize the application of transformer and its streamable variant, Emformer based acoustic model [1] for large scale speech recognition applications.
- 1358, TITLE: Emformer: Efficient Memory Transformer Based Acoustic Model for Low Latency Streaming Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9414560>  
AUTHORS: Y. Shi et al.  
HIGHLIGHT: This paper proposes an efficient memory transformer Emformer for low latency streaming speech recognition.
- 1359, TITLE: Learned Transferable Architectures Can Surpass Hand-Designed Architectures for Large Scale Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9414954>  
AUTHORS: L. He, D. Su and D. Yu  
HIGHLIGHT: In this paper, we explore the neural architecture search (NAS) for automatic speech recognition (ASR) systems.
- 1360, TITLE: Multitask Learning and Joint Optimization for Transformer-RNN-Transducer Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9414911>  
AUTHORS: J. -J. Jeon and E. Kim  
HIGHLIGHT: In this paper, we propose novel multitask learning, joint optimization, and joint decoding methods for transformer-RNN-transducer systems.
- 1361, TITLE: SEP-28k: A Dataset for Stuttering Event Detection from Podcasts with People Who Stutter  
<https://doi.org/10.1109/ICASSP39728.2021.9413520>  
AUTHORS: C. Lea, V. Mitra, A. Joshi, S. Kajarekar and J. P. Bigham  
HIGHLIGHT: In this work, we introduce Stuttering Events in Podcasts (SEP-28k), a dataset containing over 28k clips labeled with five event types including blocks, prolongations, sound repetitions, word repetitions, and interjections.
- 1362, TITLE: A Hybrid CNN-BiLSTM Voice Activity Detector  
<https://doi.org/10.1109/ICASSP39728.2021.9415081>  
AUTHORS: N. Wilkinson and T. Niesler  
HIGHLIGHT: This paper presents a new hybrid architecture for voice activity detection (VAD) incorporating both convolutional neural network (CNN) and bidirectional long short-term memory (BiLSTM) layers trained in an end-to-end manner.
- 1363, TITLE: Self-Attentive VAD: Context-Aware Detection of Voice from Noise  
<https://doi.org/10.1109/ICASSP39728.2021.9413961>  
AUTHORS: Y. R. Jo, Y. Ki Moon, W. I. Cho and G. Sik Jo

**HIGHLIGHT:** To cope with this issue with the self-attention mechanism and achieve a simple, powerful, and environment-robust VAD, we first adopt the self-attention architecture in building up the modules for voice detection and boosted prediction.

1364, **TITLE:** Preventing Early Endpointing for Online Automatic Speech Recognition

<https://doi.org/10.1109/ICASSP39728.2021.9413613>

**AUTHORS:** Y. Zhao, C. Ni, C. -C. Leung, S. Joty, E. S. Chng and B. Ma

**HIGHLIGHT:** In this paper, we propose to address the early endpointing problem from the gradient perspective.

1365, **TITLE:** MarbleNet: Deep 1D Time-Channel Separable Convolutional Neural Network for Voice Activity Detection

<https://doi.org/10.1109/ICASSP39728.2021.9414470>

**AUTHORS:** F. Jia, S. Majumdar and B. Ginsburg

**HIGHLIGHT:** We present MarbleNet, an end-to-end neural network for Voice Activity Detection (VAD).

1366, **TITLE:** Speech Enhancement Aided End-To-End Multi-Task Learning for Voice Activity Detection

<https://doi.org/10.1109/ICASSP39728.2021.9414445>

**AUTHORS:** X. Tan and X. -L. Zhang

**HIGHLIGHT:** To address this issue, here we propose a speech enhancement aided end-to-end multi-task model for VAD.

1367, **TITLE:** Robust Voice Activity Detection Using a Masked Auditory Encoder Based Convolutional Neural Network

<https://doi.org/10.1109/ICASSP39728.2021.9415045>

**AUTHORS:** N. Li et al.

**HIGHLIGHT:** Here, we propose a robust VAD approach using a masked auditory encoder based convolutional neural network (M-AECNN).

1368, **TITLE:** A Stage Match for Query-by-Example Spoken Term Detection Based On Structure Information of Query

<https://doi.org/10.1109/ICASSP39728.2021.9413442>

**AUTHORS:** J. Zhan, Q. He, J. Su and Y. Li

**HIGHLIGHT:** In this paper, we propose a stage match strategy based on the structure information of the query, represented with the unvoiced-voiced attribute of the portions in itself.

1369, **TITLE:** Knowledge Transfer for Efficient on-Device False Trigger Mitigation

<https://doi.org/10.1109/ICASSP39728.2021.9414771>

**AUTHORS:** D. Dighe, E. Marchi, S. Vishnubhotla, S. Kajarekar and D. Naik

**HIGHLIGHT:** In this paper, we address the task of determining whether a given utterance is directed towards a voice-enabled smart-assistant device or not.

1370, **TITLE:** Progressive Voice Trigger Detection: Accuracy vs Latency

<https://doi.org/10.1109/ICASSP39728.2021.9414218>

**AUTHORS:** S. Sigtia, J. Bridle, H. Richards, P. Clark, E. Marchi and V. Garg

**HIGHLIGHT:** The main idea in this work is to exploit information in words that immediately follow the trigger phrase.

1371, **TITLE:** Dynamic Curriculum Learning via Data Parameters for Noise Robust Keyword Spotting

<https://doi.org/10.1109/ICASSP39728.2021.9414501>

**AUTHORS:** T. Higuchi, S. Saxena, M. Souden, T. D. Tran, M. Delfarah and C. Dhir

**HIGHLIGHT:** Similarly, in this paper, we propose using this curriculum learning approach for acoustic modeling, and train an acoustic model on clean and noisy utterances with the data parameters.

1372, **TITLE:** CNN-Based Spoken Term Detection and Localization without Dynamic Programming

<https://doi.org/10.1109/ICASSP39728.2021.9414577>

**AUTHORS:** T. S. Fuchs, Y. Segal and J. Keshet

**HIGHLIGHT:** In this paper, we propose a spoken term detection algorithm for simultaneous prediction and localization of in-vocabulary and out-of-vocabulary terms within an audio segment.

1373, **TITLE:** Query-By-Example Keyword Spotting System Using Multi-Head Attention and Soft-triple Loss

<https://doi.org/10.1109/ICASSP39728.2021.9414156>

**AUTHORS:** J. Huang, W. Gharbieh, H. S. Shim and E. Kim

**HIGHLIGHT:** This paper proposes a neural network architecture for tackling the query-by-example user-defined keyword spotting task.

- 1374, TITLE: A Closer Look at Audio-Visual Multi-Person Speech Recognition and Active Speaker Selection  
<https://doi.org/10.1109/ICASSP39728.2021.9414160>  
AUTHORS: O. Braga and O. Siohan  
HIGHLIGHT: In the present work we further investigate this connection and examine the interplay between the two problems.
- 1375, TITLE: Generalized Knowledge Distillation from an Ensemble of Specialized Teachers Leveraging Unsupervised Neural Clustering  
<https://doi.org/10.1109/ICASSP39728.2021.9413762>  
AUTHORS: T. Fukuda and G. Kurata  
HIGHLIGHT: This paper proposes an improved generalized knowledge distillation framework with multiple dissimilar teacher networks, each of which is specialized for a specific domain, to make a deployable student network more robust to challenging acoustic environments.
- 1376, TITLE: Multistream CNN for Robust Acoustic Modeling  
<https://doi.org/10.1109/ICASSP39728.2021.9414639>  
AUTHORS: K. J. Han, J. Pan, V. K. N. Tadala, T. Ma and D. Povey  
HIGHLIGHT: This paper proposes multistream CNN, a novel neural network architecture for robust acoustic modeling in speech recognition tasks.
- 1377, TITLE: Improved Robustness to Disfluencies in Rnn-Transducer Based Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9413618>  
AUTHORS: V. Mendelev, T. Raissi, G. Camporese and M. Giollo  
HIGHLIGHT: We investigate data selection and preparation choices aiming for improved robustness of RNN-T ASR to speech disfluencies with a focus on partial words.
- 1378, TITLE: Representation Learning for Speech Recognition Using Feedback Based Relevance Weighting  
<https://doi.org/10.1109/ICASSP39728.2021.9414649>  
AUTHORS: P. Agrawal and S. Ganapathy  
HIGHLIGHT: In this work, we propose an acoustic embedding based approach for representation learning in speech recognition.
- 1379, TITLE: Towards Data Selection on TTS Data for Children's Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9413930>  
AUTHORS: W. Wang, Z. Zhou, Y. Lu, H. Wang, C. Du and Y. Qian  
HIGHLIGHT: In this work, we adopt text-to-speech data augmentation to improve the performance of children's speech recognition system.
- 1380, TITLE: An Investigation of End-to-End Models for Robust Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9414027>  
AUTHORS: A. Prasad, P. Jyothi and R. Velmurugan  
HIGHLIGHT: We address this gap and present a detailed comparison of speech enhancement-based techniques and three different model-based adaptation techniques covering data augmentation, multi-task learning, and adversarial learning for robust ASR.
- 1381, TITLE: End-to-End Dereverberation, Beamforming, and Speech Recognition with Improved Numerical Stability and Advanced Frontend  
<https://doi.org/10.1109/ICASSP39728.2021.9414464>  
AUTHORS: W. Zhang et al.  
HIGHLIGHT: In this work, we focus on the multichannel multi-speaker reverberant condition, and propose to extend our previous framework for end-to-end dereverberation, beamforming, and speech recognition with improved numerical stability and advanced frontend subnetworks including voice activity detection like masks.
- 1382, TITLE: Streaming Multi-Speaker ASR with RNN-T  
<https://doi.org/10.1109/ICASSP39728.2021.9413471>  
AUTHORS: I. Sklyar, A. Piunova and Y. Liu  
HIGHLIGHT: We investigate two approaches to multi-speaker model training of the RNN-T: deterministic output-target assignment and permutation invariant training.
- 1383, TITLE: Improving RNN Transducer with Target Speaker Extraction and Neural Uncertainty Estimation  
<https://doi.org/10.1109/ICASSP39728.2021.9414955>  
AUTHORS: J. Shi, C. Zhang, C. Weng, S. Watanabe, M. Yu and D. Yu

**HIGHLIGHT:** This work presents a joint framework that combines time-domain target-speaker speech extraction and Recurrent Neural Network Transducer (RNN-T).

1384, **TITLE:** A Progressive Learning Approach to Adaptive Noise and Speech Estimation for Speech Enhancement and Noisy Speech Recognition

<https://doi.org/10.1109/ICASSP39728.2021.9413395>

**AUTHORS:** Z. Nian, Y. -H. Tu, J. Du and C. -H. Lee

**HIGHLIGHT:** In this paper, we propose a progressive learning-based adaptive noise and speech estimation (PL-ANSE) method for speech preprocessing in noisy speech recognition, leveraging upon a frame-level noise tracking capability of improved minima controlled recursive averaging (IMCRA) and an utterance-level deep progressive learning of nonlinear interactions between speech and noise.

1385, **TITLE:** The Accented English Speech Recognition Challenge 2020: Open Datasets, Tracks, Baselines, Results and Methods

<https://doi.org/10.1109/ICASSP39728.2021.9413386>

**AUTHORS:** X. Shi et al.

**HIGHLIGHT:** The Accented English Speech Recognition Challenge (AESRC2020) is designed for providing a common testbed and promoting accent-related research.

1386, **TITLE:** Comparative Study of Different Epoch Extraction Methods for Speech Associated with Voice Disorders

<https://doi.org/10.1109/ICASSP39728.2021.9413518>

**AUTHORS:** P. Barche, K. Gurugubelli and A. K. Vuppala

**HIGHLIGHT:** Therefore, this study aimed to compare the various algorithms for detecting epoch locations from the speech associated with voice disorders.

1387, **TITLE:** Have You Made a Decision? Where? A Pilot Study on Interpretability of Polarity Analysis Based on Advising Problem

<https://doi.org/10.1109/ICASSP39728.2021.9413654>

**AUTHORS:** T. Li et al.

**HIGHLIGHT:** In this paper, we study the problem of adding interpretability to the overall polarity by predicting the meta-polarity at the same time.

1388, **TITLE:** Transformer Based Unsupervised Pre-Training for Acoustic Representation Learning

<https://doi.org/10.1109/ICASSP39728.2021.9414996>

**AUTHORS:** R. Zhang, H. Wu, W. Li, D. Jiang, W. Zou and X. Li

**HIGHLIGHT:** To handle this problem, we propose an unsupervised pre-training method using Transformer based encoder to learn a general and robust high-level representation for all acoustic tasks.

1389, **TITLE:** A Comparison of Convolutional Neural Networks for Glottal Closure Instant Detection from Raw Speech

<https://doi.org/10.1109/ICASSP39728.2021.9413675>

**AUTHORS:** J. Matoušek and D. Tihelka

**HIGHLIGHT:** In this paper, we continue to investigate the use of machine learning for the automatic detection of glottal closure instants (GCIs) from raw speech.

1390, **TITLE:** Encoder-Decoder Based Pitch Tracking and Joint Model Training for Mandarin Tone Classification

<https://doi.org/10.1109/ICASSP39728.2021.9413888>

**AUTHORS:** H. Huang, K. Wang, Y. Hu and S. Li

**HIGHLIGHT:** We propose RNN based Encoder-Decoder framework with gating mechanism which underlying models both the state cost estimation and Viterbi back-tracing pass implemented in the RAPT algorithm.

1391, **TITLE:** Construction of a Large-Scale Japanese ASR Corpus on TV Recordings

<https://doi.org/10.1109/ICASSP39728.2021.9413425>

**AUTHORS:** S. Ando and H. Fujihara

**HIGHLIGHT:** This paper presents a new large-scale Japanese speech corpus for training automatic speech recognition (ASR) systems.

1392, **TITLE:** NISP: A Multi-lingual Multi-accent Dataset for Speaker Profiling

<https://doi.org/10.1109/ICASSP39728.2021.9414349>

**AUTHORS:** S. B. Kalluri, D. Vijayaseenan, S. Ganapathy, R. R. M and P. Krishnan

**HIGHLIGHT:** In this paper, we attempt to overcome this limitation by developing a new dataset which has speech data from five different Indian languages along with English.



- 1393, TITLE: Multilingual Phonetic Dataset for Low Resource Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9413720>  
AUTHORS: X. Li, D. R. Mortensen, F. Metze and A. W. Black  
HIGHLIGHT: In this work, we present a large multilingual phonetic dataset, which is preprocessed and aligned from the UCLA phonetic dataset.
- 1394, TITLE: Age-VOX-Celeb: Multi-Modal Corpus for Facial and Speech Estimation  
<https://doi.org/10.1109/ICASSP39728.2021.9414272>  
AUTHORS: N. Tawara, A. Ogawa, Y. Kitagishi and H. Kamiyama  
HIGHLIGHT: To tackle this problem, we construct a new audio-visual age corpus named AgeVoxCeleb by annotating age labels to VoxCeleb2 videos. AgeVoxCeleb is the first large-scale, balanced, and multi-modal age corpus that contains both video and speech of the same speakers from a wide age range.
- 1395, TITLE: Didispeech: A Large Scale Mandarin Speech Corpus  
<https://doi.org/10.1109/ICASSP39728.2021.9414423>  
AUTHORS: T. Guo et al.  
HIGHLIGHT: This paper introduces a new open-sourced Mandarin speech corpus, called DiDiSpeech.
- 1396, TITLE: The in-the-Wild Speech Medical Corpus  
<https://doi.org/10.1109/ICASSP39728.2021.9414230>  
AUTHORS: J. Correia, F. Teixeira, C. Botelho, I. Trancoso and B. Raj  
HIGHLIGHT: In this work, we present the in-the-Wild Speech Medical (WSM) Corpus, a collection of in-the-wild videos, featuring subjects potentially affected by a SA disease - specifically, depression or Parkinson's disease.
- 1397, TITLE: Multiple-Hypothesis CTC-Based Semi-Supervised Adaptation of End-to-End Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9414414>  
AUTHORS: C. -T. Do, R. Doddipatla and T. Hain  
HIGHLIGHT: This paper proposes an adaptation method for end-to-end speech recognition.
- 1398, TITLE: Vowel Non-Vowel Based Spectral Warping and Time Scale Modification for Improvement in Children's ASR  
<https://doi.org/10.1109/ICASSP39728.2021.9414116>  
AUTHORS: H. Kathania, A. Kumar and M. Kurimo  
HIGHLIGHT: In this paper, we proposed a linear prediction based spectral warping method by using the knowledge of vowel and non-vowel regions in speech signals to mitigate the formant frequencies differences between child and adult speakers.
- 1399, TITLE: Extending Parrottron: An End-to-End, Speech Conversion and Speech Recognition Model for Atypical Speech  
<https://doi.org/10.1109/ICASSP39728.2021.9414644>  
AUTHORS: R. Doshi et al.  
HIGHLIGHT: We present an extended Parrottron model: a single, end-to-end network that enables voice conversion and recognition simultaneously.
- 1400, TITLE: Fundamental Frequency Feature Normalization and Data Augmentation for Child Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9413801>  
AUTHORS: G. Yeung, R. Fan and A. Alwan  
HIGHLIGHT: This study proposes a novel technique for child ASR using both feature normalization and data augmentation methods based on the relationship between formants and fundamental frequency (fo).
- 1401, TITLE: Analysis of X-Vectors for Low-Resource Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9414725>  
AUTHORS: M. Karaffa et al.  
HIGHLIGHT: The paper presents a study of usability of x-vectors for adaptation of automatic speech recognition (ASR) systems.
- 1402, TITLE: Refining Automatic Speech Recognition System for Older Adults  
<https://doi.org/10.1109/ICASSP39728.2021.9414207>  
AUTHORS: L. Chen and M. Asgari  
HIGHLIGHT: With 12 hours of training data, we attempt to develop an ASR system for socially isolated seniors (80+ years old) with possible cognitive impairments.

- 1403, TITLE: MixSpeech: Data Augmentation for Low-Resource Automatic Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9414483>  
AUTHORS: L. Meng, J. Xu, X. Tan, J. Wang, T. Qin and B. Xu  
HIGHLIGHT: In this paper, we propose MixSpeech, a simple yet effective data augmentation method based on mixup for automatic speech recognition (ASR).
- 1404, TITLE: End-to-End Multilingual Automatic Speech Recognition for Less-Resourced Languages: The Case of Four Ethiopian Languages  
<https://doi.org/10.1109/ICASSP39728.2021.9415020>  
AUTHORS: S. T. Abate, M. Y. Tachbelie and T. Schultz  
HIGHLIGHT: We have, therefore, conducted ML E2E ASR experiments for four less-resourced Ethiopian languages using different language and acoustic modelling units.
- 1405, TITLE: Improved Data Selection for Domain Adaptation in ASR  
<https://doi.org/10.1109/ICASSP39728.2021.9413869>  
AUTHORS: S. Wotherspoon, W. Hartmann, M. Snover and O. Kimball  
HIGHLIGHT: We address the problem of domain adaptation with semi-supervised training (SST).
- 1406, TITLE: Bi-APC: Bidirectional Autoregressive Predictive Coding for Unsupervised Pre-Training and its Application to Children's ASR  
<https://doi.org/10.1109/ICASSP39728.2021.9414970>  
AUTHORS: R. Fan, A. Afshan and A. Alwan  
HIGHLIGHT: We present a bidirectional unsupervised model pre-training (UPT) method and apply it to children's automatic speech recognition (ASR).
- 1407, TITLE: Meta-Adapter: Efficient Cross-Lingual Adaptation With Meta-Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9414959>  
AUTHORS: W. Hou, Y. Wang, S. Gao and T. Shinozaki  
HIGHLIGHT: In this paper, we propose to combine the adapter module with meta-learning algorithms to achieve high recognition performance under low-resource settings and improve the parameter-efficiency of the model.
- 1408, TITLE: Error-Driven Fixed-Budget ASR Personalization for Accented Speakers  
<https://doi.org/10.1109/ICASSP39728.2021.9414830>  
AUTHORS: A. Awasthi, A. Kansal, S. Sarawagi and P. Jyothi  
HIGHLIGHT: Given a speaker and an ASR model, we propose a method of identifying sentences for which the speaker's utterances are likely to be harder for the given ASR model to recognize.
- 1409, TITLE: Context-Aware Prosody Correction for Text-Based Speech Editing  
<https://doi.org/10.1109/ICASSP39728.2021.9414633>  
AUTHORS: M. Morrison, L. Rencker, Z. Jin, N. J. Bryan, J. -P. Caceres and B. Pardo  
HIGHLIGHT: In our work, we propose a new context-aware method for more natural sounding text-based editing of speech.
- 1410, TITLE: Fast DCTTS: Efficient Deep Convolutional Text-to-Speech  
<https://doi.org/10.1109/ICASSP39728.2021.9413373>  
AUTHORS: M. Kang, J. Lee, S. Kim and I. Kim  
HIGHLIGHT: We propose an end-to-end speech synthesizer, Fast DCTTS, that synthesizes speech in real time on a single CPU thread.
- 1411, TITLE: Speech Prediction in Silent Videos Using Variational Autoencoders  
<https://doi.org/10.1109/ICASSP39728.2021.9414040>  
AUTHORS: R. Yadav, A. Sardana, V. P. Namboodiri and R. M. Hegde  
HIGHLIGHT: In this paper, we present a stochastic model for generating speech in a silent video.
- 1412, TITLE: Learning Disentangled Phone and Speaker Representations in a Semi-Supervised VQ-VAE Paradigm  
<https://doi.org/10.1109/ICASSP39728.2021.9413543>  
AUTHORS: J. Williams, Y. Zhao, E. Cooper and J. Yamagishi  
HIGHLIGHT: We present a new approach to disentangle speaker voice and phone content by introducing new components to the VQ-VAE architecture for speech synthesis.

1413, TITLE: High-Intelligibility Speech Synthesis for Dysarthric Speakers with LPCNet-Based TTS and CycleVAE-Based VC  
<https://doi.org/10.1109/ICASSP39728.2021.9414136>  
AUTHORS: K. Matsubara et al.  
HIGHLIGHT: This paper presents a high-intelligibility speech synthesis method for persons with dysarthria caused by athetoid cerebral palsy.

1414, TITLE: Denoispeech: Denoising Text to Speech with Frame-Level Noise Modeling  
<https://doi.org/10.1109/ICASSP39728.2021.9413934>  
AUTHORS: C. Zhang et al.  
HIGHLIGHT: In this paper, we develop DenoiSpeech, a TTS system that can synthesize clean speech for a speaker with noisy speech data.

1415, TITLE: Non-Autoregressive Sequence-To-Sequence Voice Conversion  
<https://doi.org/10.1109/ICASSP39728.2021.9413973>  
AUTHORS: T. Hayashi, W. -C. Huang, K. Kobayashi and T. Toda  
HIGHLIGHT: This paper proposes a novel voice conversion (VC) method based on non-autoregressive sequence-to-sequence (NAR-S2S) models.

1416, TITLE: PPG-Based Singing Voice Conversion with Adversarial Representation Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9414137>  
AUTHORS: Z. Li et al.  
HIGHLIGHT: On top of recent voice conversion works, we propose a novel model to steadily convert songs while keeping their naturalness and intonation.

1417, TITLE: Litesing: Towards Fast, Lightweight and Expressive Singing Voice Synthesis  
<https://doi.org/10.1109/ICASSP39728.2021.9414043>  
AUTHORS: X. Zhuang, T. Jiang, S. -Y. Chou, B. Wu, P. Hu and S. Lui  
HIGHLIGHT: LiteSing proposed in this paper is a high-quality singing voice synthesis (SVS) system, which is fast, lightweight and expressive.

1418, TITLE: Semi-Supervised Learning for Singing Synthesis Timbre  
<https://doi.org/10.1109/ICASSP39728.2021.9413400>  
AUTHORS: J. Bonada and M. Blaauw  
HIGHLIGHT: We propose a semi-supervised singing synthesizer, which is able to learn new voices from audio data only, without any annotations such as phonetic segmentation.

1419, TITLE: Recurrent Phase Reconstruction Using Estimated Phase Derivatives from Deep Neural Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9413722>  
AUTHORS: L. Thieling, D. Wilhelm and P. Jax  
HIGHLIGHT: In this paper, we propose three changes for such a two-stage phase reconstruction system.

1420, TITLE: Stable Checkpoint Selection and Evaluation in Sequence to Sequence Speech Synthesis  
<https://doi.org/10.1109/ICASSP39728.2021.9414402>  
AUTHORS: S. Shechtman, D. Haws and R. Fernandez  
HIGHLIGHT: In this work we propose a novel stability metric designed for automatic checkpoint selection based on incomplete utterance counts within a validation set.

1421, TITLE: TSTNN: Two-Stage Transformer Based Neural Network for Speech Enhancement in the Time Domain  
<https://doi.org/10.1109/ICASSP39728.2021.9413740>  
AUTHORS: K. Wang, B. He and W. -P. Zhu  
HIGHLIGHT: In this paper, we propose a transformer-based architecture, called two-stage transformer neural network (TSTNN) for end-to-end speech denoising in the time domain.

1422, TITLE: Self-Attention Generative Adversarial Network for Speech Enhancement  
<https://doi.org/10.1109/ICASSP39728.2021.9414265>  
AUTHORS: H. Phan et al.  
HIGHLIGHT: To remedy this issue, we propose a self-attention layer adapted from non-local attention, coupled with the convolutional and deconvolutional layers of a speech enhancement GAN (SEGAN) using raw signal input.

1423, TITLE: Neural Kalman Filtering for Speech Enhancement  
<https://doi.org/10.1109/ICASSP39728.2021.9413499>  
AUTHORS: W. Xue, G. Quan, C. Zhang, G. Ding, X. He and B. Zhou  
HIGHLIGHT: In this paper, we extend the conventional Kalman filtering (KF) and propose a supervised-learning based neural Kalman filter (NKF) for speech enhancement.

1424, TITLE: Neural Noise Embedding for End-To-End Speech Enhancement with Conditional Layer Normalization  
<https://doi.org/10.1109/ICASSP39728.2021.9413931>  
AUTHORS: Z. Zhang, X. Li, Y. Li, Y. Dong, D. Wang and S. Xiong  
HIGHLIGHT: In this study, a new normalization method, termed conditional layer normalization (CLN), is introduced to improve the generalization of deep learning based speech enhancement approaches for unseen environments.

1425, TITLE: Perceptual Loss Based Speech Denoising with an Ensemble of Audio Pattern Recognition and Self-Supervised Models  
<https://doi.org/10.1109/ICASSP39728.2021.9413555>  
AUTHORS: S. Kataria, J. Villalba and N. Dehak  
HIGHLIGHT: We introduce a generalized framework called Perceptual Ensemble Regularization Loss (PERL) built on the idea of perceptual losses.

1426, TITLE: Towards An ASR Approach Using Acoustic and Language Models for Speech Enhancement  
<https://doi.org/10.1109/ICASSP39728.2021.9414565>  
AUTHORS: K. M. Nayem and D. S. Williamson  
HIGHLIGHT: In this work, we propose to modify the speech estimation process, by treating speech enhancement as a classification problem in an ASR-style manner.

1427, TITLE: A Neural Acoustic Echo Canceller Optimized Using An Automatic Speech Recognizer and Large Scale Synthetic Data  
<https://doi.org/10.1109/ICASSP39728.2021.9413585>  
AUTHORS: N. Howard, A. Park, T. Z. Shabestary, A. Gruenstein and R. Prabhavalkar  
HIGHLIGHT: Since our goal is to recognize the input speech, we consider enhancements which improve word error rates (WERs) when the predicted speech signal is passed to an automatic speech recognition (ASR) model.

1428, TITLE: Low-Complexity, Real-Time Joint Neural Echo Control and Speech Enhancement Based On Percepnet  
<https://doi.org/10.1109/ICASSP39728.2021.9414140>  
AUTHORS: J. -M. Valin, S. Tenneti, K. Helwani, U. Isik and A. Krishnaswamy  
HIGHLIGHT: In this work, we propose a system combining a traditional acoustic echo canceller, and a low-complexity joint residual echo and noise suppressor based on a hybrid signal processing/deep neural network (DSP/DNN) approach.

1429, TITLE: Acoustic Echo Cancellation with the Dual-Signal Transformation LSTM Network  
<https://doi.org/10.1109/ICASSP39728.2021.9413510>  
AUTHORS: N. L. Westhausen and B. T. Meyer  
HIGHLIGHT: This paper applies the dual-signal transformation LSTM network (DTLN) to the task of real-time acoustic echo cancellation (AEC).

1430, TITLE: High Fidelity Speech Regeneration with Application to Speech Enhancement  
<https://doi.org/10.1109/ICASSP39728.2021.9414853>  
AUTHORS: A. Polyak, L. Wolf, Y. Adi, O. Kabeli and Y. Taigman  
HIGHLIGHT: We propose a wav-to-wav generative model for speech that can generate 24khz speech in a real-time manner and which utilizes a compact speech representation, composed of ASR and identity features, to achieve a higher level of intelligibility.

1431, TITLE: A Time-Domain Convolutional Recurrent Network for Packet Loss Concealment  
<https://doi.org/10.1109/ICASSP39728.2021.9413595>  
AUTHORS: J. Lin, Y. Wang, K. Kalgaonkar, G. Keren, D. Zhang and C. Fuegen  
HIGHLIGHT: In this paper, we investigate a time-domain convolutional recurrent network (CRN) for online packet loss concealment.

1432, TITLE: Cascaded Time + Time-Frequency Unet For Speech Enhancement: Jointly Addressing Clipping, Codec Distortions, And Gaps  
<https://doi.org/10.1109/ICASSP39728.2021.9414721>  
AUTHORS: A. A. Nair and K. Koishida

**HIGHLIGHT:** In this work, we first systematically study and achieve state of the art results on each of these three distortions individually. Next, we demonstrate a neural network pipeline that cascades a time domain convolutional neural network with a time-frequency domain convolutional neural network to address all three distortions jointly.

1433, **TITLE:** Hidden Markov Model Diarisation with Speaker Location Information

<https://doi.org/10.1109/ICASSP39728.2021.9413761>

**AUTHORS:** J. H. M. Wong, X. Xiao and Y. Gong

**HIGHLIGHT:** This report proposes to extend the Hidden Markov Model (HMM) clustering method, to enable the use of speaker location information.

1434, **TITLE:** Compositional Embedding Models for Speaker Identification and Diarization with Simultaneous Speech From 2+ Speakers

<https://doi.org/10.1109/ICASSP39728.2021.9413752>

**AUTHORS:** Z. Li and J. Whitehill

**HIGHLIGHT:** We propose a new method for speaker diarization that can handle overlapping speech with 2+ people.

1435, **TITLE:** Content-Aware Speaker Embeddings for Speaker Diarisation

<https://doi.org/10.1109/ICASSP39728.2021.9414390>

**AUTHORS:** G. Sun, D. Liu, C. Zhang and P. C. Woodland

**HIGHLIGHT:** In this paper, the content-aware speaker embeddings (CASE) approach is proposed, which extends the input of the speaker classifier to include not only acoustic features but also their corresponding speech content, via phone, character, and word embeddings.

1436, **TITLE:** Multi-Scale Speaker Diarization with Neural Affinity Score Fusion

<https://doi.org/10.1109/ICASSP39728.2021.9414578>

**AUTHORS:** T. J. Park, M. Kumar and S. Narayanan

**HIGHLIGHT:** In this paper, we propose an unconventional method that tackles the trade-off between temporal resolution and the quality of the speaker representations.

1437, **TITLE:** A Comparison Study on Infant-Parent Voice Diarization

<https://doi.org/10.1109/ICASSP39728.2021.9413538>

**AUTHORS:** J. Zhu, M. Hasegawa-Johnson and N. L. McElwain

**HIGHLIGHT:** We design a framework for studying prelinguistic child voice from 3 to 24 months based on state-of-the-art algorithms in diarization.

1438, **TITLE:** End-To-End Diarization for Variable Number of Speakers with Local-Global Networks and Discriminative Speaker Embeddings

<https://doi.org/10.1109/ICASSP39728.2021.9414841>

**AUTHORS:** S. Maiti, H. Erdogan, K. Wilson, S. Wisdom, S. Watanabe and J. R. Hershey

**HIGHLIGHT:** We present an end-to-end deep network model that performs meeting diarization from single-channel audio recordings.

1439, **TITLE:** End-To-End Speaker Diarization as Post-Processing

<https://doi.org/10.1109/ICASSP39728.2021.9413436>

**AUTHORS:** S. Horiguchi, P. Garcia, Y. Fujita, S. Watanabe and K. Nagamatsu

**HIGHLIGHT:** To compensate for each other's weakness, we propose to use a two-speaker end-to-end diarization method as post-processing of the results obtained by a clustering-based method.

1440, **TITLE:** BW-EDA-EEND: streaming END-TO-END Neural Speaker Diarization for a Variable Number of Speakers

<https://doi.org/10.1109/ICASSP39728.2021.9414371>

**AUTHORS:** E. Han, C. Lee and A. Stolcke

**HIGHLIGHT:** We present a novel online end-to-end neural diarization system, BW-EDA-EEND, that processes data incrementally for a variable number of speakers.

1441, **TITLE:** Integrating End-to-End Neural and Clustering-Based Diarization: Getting the Best of Both Worlds

<https://doi.org/10.1109/ICASSP39728.2021.9414333>

**AUTHORS:** K. Kinoshita, M. Delcroix and N. Tawara

**HIGHLIGHT:** In this paper, we propose a simple but effective hybrid diarization framework that works with overlapped speech and for long recordings containing an arbitrary number of speakers.

- 1442, TITLE: Siamese Capsule Network for End-to-End Speaker Recognition in the Wild  
<https://doi.org/10.1109/ICASSP39728.2021.9414722>  
AUTHORS: A. Hajavi and A. Etemad  
HIGHLIGHT: We propose an end-to-end deep model for speaker verification in the wild.
- 1443, TITLE: A Real-Time Speaker Diarization System Based on Spatial Spectrum  
<https://doi.org/10.1109/ICASSP39728.2021.9413544>  
AUTHORS: S. Zheng, W. Huang, X. Wang, H. Suo, J. Feng and Z. Yan  
HIGHLIGHT: In this paper we describe a speaker diarization system that enables localization and identification of all speakers present in a conversation or meeting.
- 1444, TITLE: Unsupervised Neural Adaptation Model Based on Optimal Transport for Spoken Language Identification  
<https://doi.org/10.1109/ICASSP39728.2021.9414045>  
AUTHORS: X. Lu, P. Shen, Y. Tsao and H. Kawai  
HIGHLIGHT: In this paper, we propose an unsupervised neural adaptation model to deal with the distribution mismatch problem for SLID.
- 1445, TITLE: Joint ASR and Language Identification Using RNN-T: An Efficient Approach to Dynamic Language Switching  
<https://doi.org/10.1109/ICASSP39728.2021.9413734>  
AUTHORS: S. Punjabi et al.  
HIGHLIGHT: Since this solution is neither scalable nor cost- and memory-efficient, especially for on-device applications, we propose end-to-end, streaming, joint ASR-LID architectures based on the recurrent neural network transducer framework.
- 1446, TITLE: Spoken Language Identification in Unseen Target Domain Using Within-Sample Similarity Loss  
<https://doi.org/10.1109/ICASSP39728.2021.9414090>  
AUTHORS: M. H. S. Kapoor, D. A. Dinesh and P. Rajan  
HIGHLIGHT: To address this, we propose an auxiliary within-sample similarity loss (WSSL) which encourages the network to suppress the channel-specific contents in the speech.
- 1447, TITLE: Exploring the use of Common Label Set to Improve Speech Recognition of Low Resource Indian Languages  
<https://doi.org/10.1109/ICASSP39728.2021.9414961>  
AUTHORS: V. M. Shetty and S. Umesh  
HIGHLIGHT: Since the visual rendering of these characters is different, in this paper, we explore the benefits of representing such similar target subword units (e.g., Byte Pair Encoded(BPE) units) through a Common Label Set (CLS).
- 1448, TITLE: Phone Distribution Estimation for Low Resource Languages  
<https://doi.org/10.1109/ICASSP39728.2021.9415014>  
AUTHORS: X. Li, J. Li, J. Yao, A. W. Black and F. Metze  
HIGHLIGHT: In this work, we propose a novel approach to estimate phone distributions by only requiring raw audio datasets: We first estimate the phone ranks by combining language-independent recognition results and Learning to Rank results.
- 1449, TITLE: How Phonotactics Affect Multilingual and Zero-Shot ASR Performance  
<https://doi.org/10.1109/ICASSP39728.2021.9414478>  
AUTHORS: S. Feng et al.  
HIGHLIGHT: We show that the gain from modeling crosslingual phonotactics is limited, and imposing a too strong model can hurt the zero-shot transfer.
- 1450, TITLE: Modelling Paralinguistic Properties in Conversational Speech to Detect Bipolar Disorder and Borderline Personality Disorder  
<https://doi.org/10.1109/ICASSP39728.2021.9413891>  
AUTHORS: B. Wang, Y. Wu, N. Vaci, M. Liakata, T. Lyons and K. E. A. Saunders  
HIGHLIGHT: In this work, we investigate the automatic detection of these two conditions by modelling both verbal and non-verbal cues in a set of interviews.
- 1451, TITLE: An Attention Model for Hypernasality Prediction in Children with Cleft Palate  
<https://doi.org/10.1109/ICASSP39728.2021.9414860>  
AUTHORS: V. C. Mathad, N. Scherer, K. Chapman, J. Liss and V. Berisha  
HIGHLIGHT: In this work, we propose an attention-based bidirectional long-short memory (BLSTM) model that directly maps the frame-level features to utterance-level ratings by focusing only on specific speech frames carrying hyper-nasal cues.

- 1452, TITLE: An End-to-End Speech Accent Recognition Method Based on Hybrid CTC/Attention Transformer ASR  
<https://doi.org/10.1109/ICASSP39728.2021.9414082>  
AUTHORS: Q. Gao, H. Wu, Y. Sun and Y. Duan  
HIGHLIGHT: This paper proposes a novel accent recognition system in the framework of a transformer-based end-to-end speech recognition system.
- 1453, TITLE: Multi-Task Estimation of Age and Cognitive Decline from Speech  
<https://doi.org/10.1109/ICASSP39728.2021.9414642>  
AUTHORS: Y. Pan, V. S. Nallanthighal, D. Blackburn, H. Christensen and A. H?rm?  
HIGHLIGHT: In this paper, multi-task learning is applied for the joint estimation of age and the Mini-Mental Status Evaluation criteria (MMSE) commonly used to assess cognitive decline.
- 1454, TITLE: Deepemocluster: a Semi-Supervised Framework for Latent Cluster Representation of Speech Emotions  
<https://doi.org/10.1109/ICASSP39728.2021.9414035>  
AUTHORS: W. -C. Lin, K. Sridhar and C. Busso  
HIGHLIGHT: In this study, we introduce a new SSL framework, which we refer to as the DeepEmoCluster framework, for attribute-based SER tasks.
- 1455, TITLE: The Role of Task and Acoustic Similarity in Audio Transfer Learning: Insights from the Speech Emotion Recognition Case  
<https://doi.org/10.1109/ICASSP39728.2021.9414896>  
AUTHORS: A. Triantafyllopoulos and B. W. Schuller  
HIGHLIGHT: We perform a similar investigation for the case of speech emotion recognition (SER), and conclude that transfer learning for SER is influenced both by the choice of pre-training task and by the differences in acoustic conditions between the upstream and downstream data sets, with the former having a bigger impact.
- 1456, TITLE: Speech-Based Depression Prediction Using Encoder-Weight-Only Transfer Learning and a Large Corpus  
<https://doi.org/10.1109/ICASSP39728.2021.9414208>  
AUTHORS: A. Harati, E. Shriberg, T. Rutowski, P. Chlebek, Y. Lu and R. Oliveira  
HIGHLIGHT: We explore a speech-based transfer learning approach that uses a lightweight encoder and that transfers only the encoder weights, enabling a simplified run-time model.
- 1457, TITLE: Estimating Severity of Depression From Acoustic Features and Embeddings of Natural Speech  
<https://doi.org/10.1109/ICASSP39728.2021.9414129>  
AUTHORS: S. H. Dumpala, S. Rempel, K. Dikaos, M. Sajjadian, R. Uher and S. Oore  
HIGHLIGHT: In this work, we examine two novel approaches for improving depression severity estimation from short audio recordings of speech.
- 1458, TITLE: Automatic Elicitation Compliance for Short-Duration Speech Based Depression Detection  
<https://doi.org/10.1109/ICASSP39728.2021.9414366>  
AUTHORS: B. Stasak, Z. Huang, D. Joachim and J. Epps  
HIGHLIGHT: In this multi-corpus study of over 25,000 ?pataka? utterances, it was discovered that speech landmark- based features were sensitive to the number of ?pataka? utterances per recording.
- 1459, TITLE: Deep Neural Network Embeddings for the Estimation of the Degree of Sleepiness  
<https://doi.org/10.1109/ICASSP39728.2021.9413589>  
AUTHORS: J. V. Egas-L?pez and G. Gosztolya  
HIGHLIGHT: In this study, we employ the x-vector approach, currently the state-of-the-art in speaker recognition, as a neural network feature extractor to detect the level of sleepiness of a speaker.
- 1460, TITLE: Pause-Encoded Language Models for Recognition of Alzheimer?s Disease and Emotion  
<https://doi.org/10.1109/ICASSP39728.2021.9413548>  
AUTHORS: J. Yuan, X. Cai and K. Church  
HIGHLIGHT: We propose enhancing Transformer language models (BERT, RoBERTa) to take advantage of pauses.
- 1461, TITLE: End-2-End Modeling of Speech and Gait from Patients with Parkinson?s Disease: Comparison Between High Quality Vs. Smartphone Data  
<https://doi.org/10.1109/ICASSP39728.2021.9414729>  
AUTHORS: J. C. Vasquez-Correa, T. Arias-Vergara, P. Klumpp, P. A. Perez-Toro, J. R. Orozco-Arroyave and E. N?th  
HIGHLIGHT: We propose the use of state-of-the-art deep learning techniques to evaluate the speech and gait symptoms of patients.



- 1462, TITLE: A Sequential Contrastive Learning Framework for Robust Dysarthric Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9415017>  
AUTHORS: L. Wu, D. Zong, S. Sun and J. Zhao  
HIGHLIGHT: In this paper, we propose a contrastive learning framework for robust dysarthric speech recognition (DSR) by capturing the dysarthric speech variability.
- 1463, TITLE: Automatic And Perceptual Discrimination Between Dysarthria, Apraxia of Speech, and Neurotypical Speech  
<https://doi.org/10.1109/ICASSP39728.2021.9414283>  
AUTHORS: I. Kodrasi, M. Pernon, M. Laganaro and H. Bourlard  
HIGHLIGHT: In this paper, we investigate a three-class automatic technique and a set of handcrafted features for the discrimination of dysarthria, AoS and neurotypical speech.
- 1464, TITLE: Effect of Noise and Model Complexity on Detection of Amyotrophic Lateral Sclerosis and Parkinson's Disease Using Pitch and MFCC  
<https://doi.org/10.1109/ICASSP39728.2021.9413997>  
AUTHORS: T. Bhattacharjee et al.  
HIGHLIGHT: This study aims to examine the robustness of these cues against background noise and model complexity, which has not been investigated before.
- 1465, TITLE: Multi-Task Transformer with Input Feature Reconstruction for Dysarthric Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9414614>  
AUTHORS: C. Ding, S. Sun and J. Zhao  
HIGHLIGHT: To this end, we propose a multi-task Transformer with input feature reconstruction as an auxiliary task, where the main task of DSR and the auxiliary reconstruction task share the same encoder network.
- 1466, TITLE: Detecting Alzheimer's Disease from Speech Using Neural Networks with Bottleneck Features and Data Augmentation  
<https://doi.org/10.1109/ICASSP39728.2021.9413566>  
AUTHORS: Z. Liu, Z. Guo, Z. Ling and Y. Li  
HIGHLIGHT: This paper presents a method of detecting Alzheimer's disease (AD) from the spontaneous speech of subjects in a picture description task using neural networks.
- 1467, TITLE: Automatic Dysarthric Speech Detection Exploiting Pairwise Distance-Based Convolutional Neural Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9413922>  
AUTHORS: P. Janbakhshi, I. Kodrasi and H. Bourlard  
HIGHLIGHT: In this paper we propose a novel automatic dysarthric speech detection approach based on analyses of pairwise distance matrices using convolutional neural networks (CNNs).
- 1468, TITLE: Improved Neural Language Model Fusion for Streaming Recurrent Neural Network Transducer  
<https://doi.org/10.1109/ICASSP39728.2021.9414784>  
AUTHORS: S. Kim et al.  
HIGHLIGHT: In this paper, we propose extensions to these techniques that allow RNN-T to exploit external NNLMs during both training and inference time, resulting in 13-18% relative Word Error Rate improvement on Librispeech compared to strong baselines.
- 1469, TITLE: Internal Language Model Training for Domain-Adaptive End-To-End Speech Recognition  
<https://doi.org/10.1109/ICASSP39728.2021.9415039>  
AUTHORS: Z. Meng et al.  
HIGHLIGHT: To improve the ILME-based inference, we propose an internal LM training (ILMT) method to minimize an additional internal LM loss by updating only the E2E model components that affect the internal LM estimation.
- 1470, TITLE: Speech Recognition by Simply Fine-Tuning Bert  
<https://doi.org/10.1109/ICASSP39728.2021.9413668>  
AUTHORS: W. -C. Huang, C. -H. Wu, S. -B. Luo, K. -Y. Chen, H. -M. Wang and T. Toda  
HIGHLIGHT: We propose a simple method for automatic speech recognition (ASR) by fine-tuning BERT, which is a language model (LM) trained on large-scale unlabeled text data and can generate rich contextual representations.
- 1471, TITLE: Personalization Strategies for End-to-End Speech Recognition Systems  
<https://doi.org/10.1109/ICASSP39728.2021.9413962>

AUTHORS: A. Gourav et al.

HIGHLIGHT: In this work, we demonstrate how first- and second-pass rescoring strategies can be leveraged together to improve the recognition of such words.

1472, TITLE: Improving Entity Recall in Automatic Speech Recognition with Neural Embeddings

<https://doi.org/10.1109/ICASSP39728.2021.9413756>

AUTHORS: C. Li, P. Rondon, D. Caseiro, L. Velikovich, X. Velez and P. Aleksic

HIGHLIGHT: In this work, we present a method which uses learned text embeddings and nearest neighbor retrieval within a large database of entity embeddings to correct misrecognitions.

1473, TITLE: Adaptable Multi-Domain Language Model for Transformer ASR

<https://doi.org/10.1109/ICASSP39728.2021.9413475>

AUTHORS: T. Lee et al.

HIGHLIGHT: We propose an adapter based multi-domain Transformer based language model (LM) for Transformer ASR.

1474, TITLE: Transformer Language Models with LSTM-Based Cross-Utterance Information Representation

<https://doi.org/10.1109/ICASSP39728.2021.9414477>

AUTHORS: G. Sun, C. Zhang and P. C. Woodland

HIGHLIGHT: To extract more powerful and robust cross-utterance representations for the Transformer LM (TLM), this paper proposes the R-TLM which uses hidden states in a long short-term memory (LSTM) LM.

1475, TITLE: Large Margin Training Improves Language Models for ASR

<https://doi.org/10.1109/ICASSP39728.2021.9414724>

AUTHORS: J. Wang, J. Huang and K. W. Church

HIGHLIGHT: In this work, we propose a large margin language model (LMLM).

1476, TITLE: Domain-Aware Neural Language Models for Speech Recognition

<https://doi.org/10.1109/ICASSP39728.2021.9414800>

AUTHORS: L. Liu et al.

HIGHLIGHT: We present a domain-aware rescoring framework suitable for achieving domain-adaptation during second-pass rescoring in production settings.

1477, TITLE: Bayesian Transformer Language Models for Speech Recognition

<https://doi.org/10.1109/ICASSP39728.2021.9414046>

AUTHORS: B. Xue et al.

HIGHLIGHT: In order to address these issues, this paper proposes a full Bayesian learning framework for Transformer LM estimation.

1478, TITLE: Mixed Precision Quantization of Transformer Language Models for Speech Recognition

<https://doi.org/10.1109/ICASSP39728.2021.9414076>

AUTHORS: J. Xu, S. Hu, J. Yu, X. Liu and H. Meng

HIGHLIGHT: To this end, novel mixed precision DNN quantization methods are proposed in this paper.

1479, TITLE: Federated Marginal Personalization for ASR Rescoring

<https://doi.org/10.1109/ICASSP39728.2021.9415056>

AUTHORS: Z. Liu and F. Peng

HIGHLIGHT: We introduce federated marginal personalization (FMP), a novel method for continuously updating personalized neural network language models (NNLMs) on private devices using federated learning (FL).

1480, TITLE: Multi Path Training Framework for Data-Driven Open-Domain Conversation System

<https://doi.org/10.1109/ICASSP39728.2021.9414273>

AUTHORS: S. Wu, D. Zhang, Y. Li and Z. Wu

HIGHLIGHT: This paper proposes a novel framework, Multi-Path Training (MPT), for training a robust dialogue response generation system.

1481, TITLE: Action State Update Approach to Dialogue Management

<https://doi.org/10.1109/ICASSP39728.2021.9414888>

AUTHORS: S. Stoyanchev, S. Keizer and R. Doddipatla

HIGHLIGHT: We propose the action state update approach (ASU) for utterance interpretation, featuring a statistically trained binary classifier used to detect dialogue state update actions in the text of a user utterance.

- 1482, TITLE: Generating Empathetic Responses by Injecting Anticipated Emotion  
<https://doi.org/10.1109/ICASSP39728.2021.9413596>  
AUTHORS: Y. Liu, J. Du, X. Li and R. Xu  
HIGHLIGHT: In this paper, we propose a novel empathetic response generation method that incorporates the anticipated emotion into response generation by minimizing the divergence between distribution of responses? anticipated emotion and ground-truth emotion.
- 1483, TITLE: Towards Immediate Backchannel Generation Using Attention-Based Early Prediction Model  
<https://doi.org/10.1109/ICASSP39728.2021.9414193>  
AUTHORS: A. I. Adiba, T. Homma and T. Miyoshi  
HIGHLIGHT: To make accurate predictions on the basis of delayed ASR outputs, we propose early prediction for backchannel opportunity and backchannel category based on attention-based LSTM mechanisms.
- 1484, TITLE: Error-Driven Pruning of Language Models for Virtual Assistants  
<https://doi.org/10.1109/ICASSP39728.2021.9415035>  
AUTHORS: S. Gondala, L. Verwimp, E. Pusateri, M. Tsagkias and C. Van Gysel  
HIGHLIGHT: We customize entropy pruning by allowing for a keep list of infrequent n-grams that require a more relaxed pruning threshold, and propose three methods to construct the keep list.
- 1485, TITLE: Paragraph Level Multi-Perspective Context Modeling for Question Generation  
<https://doi.org/10.1109/ICASSP39728.2021.9414933>  
AUTHORS: J. Bai, W. Rong, F. Xia, Y. Wang, Y. Ouyang and Z. Xiong  
HIGHLIGHT: In this research, we proposed a multi-perspective paragraph context modeling mechanism, which firstly encodes the contextualized representation of input paragraph, and then utilize multi-head self-attention and Rezero network to further enhance paragraph-level feature extraction and context modeling.
- 1486, TITLE: Improving Dialogue Response Generation Via Knowledge Graph Filter  
<https://doi.org/10.1109/ICASSP39728.2021.9414324>  
AUTHORS: Y. Wang, Y. Wang, X. Lou, W. Rong, Z. Hao and S. Wang  
HIGHLIGHT: In this paper, we propose to leverage the contextual word representation of dialog post to filter out irrelevant knowledge with an attention-based triple filter network.
- 1487, TITLE: Topic-Aware Dialogue Generation with Two-Hop Based Graph Attention  
<https://doi.org/10.1109/ICASSP39728.2021.9414472>  
AUTHORS: S. Zhou, W. Rong, J. Zhang, Y. Wang, L. Shi and Z. Xiong  
HIGHLIGHT: For this purpose, we propose an open-domain topic-aware dialogue generation model via joint learning.
- 1488, TITLE: HSAN: A Hierarchical Self-Attention Network for Multi-Turn Dialogue Generation  
<https://doi.org/10.1109/ICASSP39728.2021.9413753>  
AUTHORS: Y. Kong, L. Zhang, C. Ma and C. Cao  
HIGHLIGHT: In this paper, we propose a hierarchical self-attention network, named HSAN, which attends to the important words and utterances in context simultaneously.
- 1489, TITLE: Learning to Select Context in a Hierarchical and Global Perspective for Open-Domain Dialogue Generation  
<https://doi.org/10.1109/ICASSP39728.2021.9414730>  
AUTHORS: L. Shen, H. Zhan, X. Shen and Y. Feng  
HIGHLIGHT: In this paper, we propose a novel model with hierarchical self-attention mechanism and distant supervision to not only detect relevant words and utterances in short and long distances, but also discern related information globally when decoding.
- 1490, TITLE: Towards Efficiently Diversifying Dialogue Generation Via Embedding Augmentation  
<https://doi.org/10.1109/ICASSP39728.2021.9414915>  
AUTHORS: Y. Cao, L. Ding, Z. Tian and M. Fang  
HIGHLIGHT: Unlike previous augmentation methods that mostly focus on token manipulation and ignore the essential variety within a single sample using hard labels, we propose to promote the generation diversity of the neural dialogue models via soft embedding augmentation along with soft labels in this paper.
- 1491, TITLE: End2End Acoustic to Semantic Transduction  
<https://doi.org/10.1109/ICASSP39728.2021.9413581>  
AUTHORS: V. Pelloin et al.

**HIGHLIGHT:** In this paper, we propose a novel end-to-end sequence-to-sequence spoken language understanding model using an attention mechanism.

1492, **TITLE:** Acoustics Based Intent Recognition Using Discovered Phonetic Units for Low Resource Languages  
<https://doi.org/10.1109/ICASSP39728.2021.9415112>

**AUTHORS:** A. Gupta, X. Li, S. K. Rallabandi and A. W. Black

**HIGHLIGHT:** With the aim of aiding development of spoken dialog systems in low resourced languages, we propose a novel acoustics based intent recognition system that uses discovered phonetic units for intent classification.

1493, **TITLE:** Speech-Language Pre-Training for End-to-End Spoken Language Understanding  
<https://doi.org/10.1109/ICASSP39728.2021.9414900>

**AUTHORS:** Y. Qian et al.

**HIGHLIGHT:** In this paper, we propose to unify a well-optimized E2E ASR encoder (speech) and a pre-trained language model encoder (language) into a transformer decoder.

1494, **TITLE:** Two-Stage Textual Knowledge Distillation for End-to-End Spoken Language Understanding  
<https://doi.org/10.1109/ICASSP39728.2021.9414619>

**AUTHORS:** S. Kim, G. Kim, S. Shin and S. Lee

**HIGHLIGHT:** To utilize textual information more effectively, this work proposes a two-stage textual knowledge distillation method that matches utterancelevel representations and predicted logits of two modalities during pre-training and fine-tuning, sequentially.

1495, **TITLE:** Semi-Supervised Spoken Language Understanding via Self-Supervised Speech and Language Model Pretraining  
<https://doi.org/10.1109/ICASSP39728.2021.9414922>

**AUTHORS:** C. -I. Lai, Y. -S. Chuang, H. -Y. Lee, S. -W. Li and J. Glass

**HIGHLIGHT:** In this paper, we propose a clean and general framework to learn semantics directly from speech with semi-supervision from transcribed or untranscribed speech to address these issues.

1496, **TITLE:** DO as I Mean, Not as I Say: Sequence Loss Training for Spoken Language Understanding  
<https://doi.org/10.1109/ICASSP39728.2021.9414566>

**AUTHORS:** M. Rao et al.

**HIGHLIGHT:** In this work, we propose non-differentiable sequence losses based on SLU metrics as a proxy for semantic error and use the REINFORCE trick to train ASR and SLU models with this loss.

1497, **TITLE:** St-Bert: Cross-Modal Language Model Pre-Training for End-to-End Spoken Language Understanding  
<https://doi.org/10.1109/ICASSP39728.2021.9414558>

**AUTHORS:** M. Kim, G. Kim, S. -W. Lee and J. -W. Ha

**HIGHLIGHT:** In this context, we introduce a cross-modal pre-trained language model, called Speech-Text BERT (ST-BERT), to tackle end-to-end spoken language understanding (E2E SLU) tasks.

1498, **TITLE:** End-to-End Spoken Language Understanding Using Transformer Networks and Self-Supervised Pre-Trained Features  
<https://doi.org/10.1109/ICASSP39728.2021.9414522>

**AUTHORS:** E. Morais, H. -K. J. Kuo, S. Thomas, Z. T'ske and B. Kingsbury

**HIGHLIGHT:** In this paper we introduce a modular End-to-End (E2E) SLU transformer network based architecture which allows the use of self-supervised pre- trained acoustic features, pre-trained model initialization and multi-task training.

1499, **TITLE:** Sentiment Injected Iteratively Co-Interactive Network for Spoken Language Understanding  
<https://doi.org/10.1109/ICASSP39728.2021.9413885>

**AUTHORS:** Z. Huang, F. Liu, P. Zhou and Y. Zou

**HIGHLIGHT:** In this paper, we argue that implicitly introducing sentimental features can promote SLU performance.

1500, **TITLE:** RNN Transducer Models for Spoken Language Understanding  
<https://doi.org/10.1109/ICASSP39728.2021.9414029>

**AUTHORS:** S. Thomas et al.

**HIGHLIGHT:** We present a comprehensive study on building and adapting RNN transducer (RNN-T) models for spoken language understanding (SLU).

- 1501, TITLE: Leveraging Acoustic and Linguistic Embeddings from Pretrained Speech and Language Models for Intent Classification  
<https://doi.org/10.1109/ICASSP39728.2021.9413388>  
AUTHORS: B. Sharma, M. Madhavi and H. Li  
HIGHLIGHT: In this work, we propose a novel intent classification framework that employs acoustic features extracted from a pretrained speech recognition system and linguistic features learned from a pretrained language model.
- 1502, TITLE: ORTHROS: non-autoregressive end-to-end speech translation With dual-decoder  
<https://doi.org/10.1109/ICASSP39728.2021.9415093>  
AUTHORS: H. Inaguma, Y. Higuchi, K. Duh, T. Kawahara and S. Watanabe  
HIGHLIGHT: We propose a novel NAR E2E-ST framework, Orthros, in which both NAR and autoregressive (AR) decoders are jointly trained on the shared speech encoder.
- 1503, TITLE: Cascaded Models with Cyclic Feedback for Direct Speech Translation  
<https://doi.org/10.1109/ICASSP39728.2021.9413719>  
AUTHORS: T. K. Lam, S. Schamoni and S. Riezler  
HIGHLIGHT: We present a technique that allows cascades of automatic speech recognition (ASR) and machine translation (MT) to exploit in-domain direct speech translation data in addition to out-of-domain MT and ASR data.
- 1504, TITLE: Jointly Trained Transformers Models for Spoken Language Translation  
<https://doi.org/10.1109/ICASSP39728.2021.9414159>  
AUTHORS: H. K. Vydan, M. Karafi?t, K. Zmolikova, L. Burget and H. Cernock?  
HIGHLIGHT: In this work, degradation in performance is reduced by creating an End-to-End differentiable pipeline between the ASR and MT systems.
- 1505, TITLE: Efficient Use of End-to-End Data in Spoken Language Processing  
<https://doi.org/10.1109/ICASSP39728.2021.9414510>  
AUTHORS: Y. Lu, Y. Wang and M. J. F. Gales  
HIGHLIGHT: This work examines how the best use can be made of limited end-to-end training for sequence-to-sequence tasks.
- 1506, TITLE: Streaming Simultaneous Speech Translation with Augmented Memory Transformer  
<https://doi.org/10.1109/ICASSP39728.2021.9414897>  
AUTHORS: X. Ma, Y. Wang, M. J. Dousti, P. Koehn and J. Pino  
HIGHLIGHT: We propose an end-to-end transformer-based sequence-to-sequence model, equipped with an augmented memory transformer encoder, which has shown great success on the streaming automatic speech recognition task with hybrid or transducer-based models.
- 1507, TITLE: An Empirical Study of End-To-End Simultaneous Speech Translation Decoding Strategies  
<https://doi.org/10.1109/ICASSP39728.2021.9414276>  
AUTHORS: H. Nguyen, Y. Est?ve and L. Besacier  
HIGHLIGHT: This paper proposes a decoding strategy for end-to-end simultaneous speech translation.
- 1508, TITLE: Modeling Homophone Noise for Robust Neural Machine Translation  
<https://doi.org/10.1109/ICASSP39728.2021.9413586>  
AUTHORS: W. Qin, X. Li, Y. Sun, D. Xiong, J. Cui and B. Wang  
HIGHLIGHT: In this paper, we propose a robust neural machine translation (NMT) framework to deal with homophone errors.
- 1509, TITLE: Machine Translation Verbosity Control for Automatic Dubbing  
<https://doi.org/10.1109/ICASSP39728.2021.9414411>  
AUTHORS: S. M. Lakew et al.  
HIGHLIGHT: In this paper, we focus on the problem of controlling the verbosity of machine translation out-put, so that subsequent steps of our automatic dubbing pipeline can generate dubs of better quality.
- 1510, TITLE: Improvements to Prosodic Alignment for Automatic Dubbing  
<https://doi.org/10.1109/ICASSP39728.2021.9414966>  
AUTHORS: Y. Virkar, M. Federico, R. Enyedi and R. Barra-Chicote  
HIGHLIGHT: To this end, we present improvements to the prosodic alignment component of our recently introduced dubbing architecture.

1511, TITLE: Image-Assisted Transformer in Zero-Resource Multi-Modal Translation  
<https://doi.org/10.1109/ICASSP39728.2021.9413389>  
AUTHORS: P. Huang, S. Sun and H. Yang  
HIGHLIGHT: In this paper, we investigate how to use visual information as an auxiliary hint for a Transformer-based system in a zero-resource translation scenario.

1512, TITLE: Sentence Boundary Augmentation for Neural Machine Translation Robustness  
<https://doi.org/10.1109/ICASSP39728.2021.9413492>  
AUTHORS: D. Li, T. I. N. Arivazhagan, C. Cherry and D. Padfield  
HIGHLIGHT: Through in-depth error analysis, we show that sentence boundary segmentation has the largest impact on quality, and we develop a simple data augmentation strategy to improve segmentation robustness.

1513, TITLE: An Empirical Study on Task-Oriented Dialogue Translation  
<https://doi.org/10.1109/ICASSP39728.2021.9413521>  
AUTHORS: S. Liu  
HIGHLIGHT: In this paper, we systematically investigate advanced models on the task-oriented dialogue translation task, including sentence-level, document-level and non-autoregressive NMT models.

1514, TITLE: MAPGN: Masked Pointer-Generator Network for Sequence-to-Sequence Pre-Training  
<https://doi.org/10.1109/ICASSP39728.2021.9414738>  
AUTHORS: M. Ihori, N. Makishima, T. Tanaka, A. Takashima, S. Orihashi and R. Masumura  
HIGHLIGHT: This paper presents a self-supervised learning method for pointer-generator networks to improve spoken-text normalization.

1515, TITLE: Aligning the training and evaluation of unsupervised text style Transfer  
<https://doi.org/10.1109/ICASSP39728.2021.9414515>  
AUTHORS: W. Qian, F. Zhu, J. Yang, J. Han and S. Hu  
HIGHLIGHT: This paper proposes a novel training method based on the evaluation metrics to address the discrepancy issue.

1516, TITLE: Neural Inverse Text Normalization  
<https://doi.org/10.1109/ICASSP39728.2021.9414912>  
AUTHORS: M. Sunkara, C. Shivade, S. Bodapati and K. Kirchhoff  
HIGHLIGHT: We propose an efficient and robust neural solution for ITN leveraging transformer based seq2seq models and FST-based text normalization techniques for data preparation.

1517, TITLE: Generating Human Readable Transcript for Automatic Speech Recognition with Pre-Trained Language Model  
<https://doi.org/10.1109/ICASSP39728.2021.9414626>  
AUTHORS: J. Liao et al.  
HIGHLIGHT: In this work, we propose an ASR post-processing model that aims to transform the incorrect and noisy ASR output into a readable text for humans and downstream tasks.

1518, TITLE: Improving Neural Text Normalization with Partial Parameter Generator and Pointer-Generator Network  
<https://doi.org/10.1109/ICASSP39728.2021.9415113>  
AUTHORS: W. Jiang, J. Li, M. Chen, J. Ma, S. Wang and J. Xiao  
HIGHLIGHT: In this paper, we treat TN as a neural machine translation problem and present a pure data-driven TN system using Transformer framework.

1519, TITLE: Incorporating Syntactic and Phonetic Information into Multimodal Word Embeddings Using Graph Convolutional Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414148>  
AUTHORS: W. ZHU, S. LIU and C. LIU  
HIGHLIGHT: This inspires us to propose a new multimodal word representation model, namely, GCNW, which uses the graph convolutional network to incorporate the phonetic and syntactic information into the word representation.

1520, TITLE: LIFI: Towards Linguistically Informed Frame Interpolation  
<https://doi.org/10.1109/ICASSP39728.2021.9413998>  
AUTHORS: A. N. Mathur, D. Batra, Y. K. Singla, R. Ratn Shah, C. Chen and R. Zimmermann  
HIGHLIGHT: Here we explore the problem of speech video interpolation.

1521, TITLE: Triple Sequence Generative Adversarial Nets for Unsupervised Image Captioning

<https://doi.org/10.1109/ICASSP39728.2021.9414335>

AUTHORS: Y. Zhou, W. Tao and W. Zhang

HIGHLIGHT: In this paper, we present a novel triple sequence generative adversarial net including an image generator, a discriminator, and a sentence generator.

1522, TITLE: Align or attend? Toward More Efficient and Accurate Spoken Word Discovery Using Speech-to-Image Retrieval

<https://doi.org/10.1109/ICASSP39728.2021.9414418>

AUTHORS: L. Wang, X. Wang, M. Hasegawa-Johnson, O. Scharenborg and N. Dehak

HIGHLIGHT: We verify our theory by conducting retrieval and word discovery experiments on MSCOCO and Flickr8k, and empirically demonstrate that both neural MT with self-attention and statistical MT achieve word discovery scores that are superior to those of a state-of-the-art neural retrieval system, outperforming it by 2% and 5% alignment F1 scores respectively.

1523, TITLE: Towards Practical Lipreading with Distilled and Efficient Models

<https://doi.org/10.1109/ICASSP39728.2021.9415063>

AUTHORS: P. Ma, B. Martinez, S. Petridis and M. Pantic

HIGHLIGHT: In this work, we propose a series of innovations that significantly bridge that gap: first, we raise the state-of-the-art performance by a wide margin on LRW and LRW-1000 to 88.5 % and 46.6 %, respectively using self-distillation.

1524, TITLE: End-To-End Audio-Visual Speech Recognition with Conformers

<https://doi.org/10.1109/ICASSP39728.2021.9414567>

AUTHORS: P. Ma, S. Petridis and M. Pantic

HIGHLIGHT: In this work, we present a hybrid CTC/Attention model based on a ResNet-18 and Convolution-augmented transformer (Conformer), that can be trained in an end-to-end manner.

1525, TITLE: ASR N-Best Fusion Nets

<https://doi.org/10.1109/ICASSP39728.2021.9414806>

AUTHORS: X. Liu et al.

HIGHLIGHT: This paper proposes a fusion network to jointly consider ASR n-best hypotheses for enhanced robustness to ASR errors.

1526, TITLE: Boosting Low-Resource Intent Detection with in-Scope Prototypical Networks

<https://doi.org/10.1109/ICASSP39728.2021.9414548>

AUTHORS: H. Lin, Y. Yan and G. Chen

HIGHLIGHT: In this paper, we propose a universal In-scope Prototypical Networks for low-resource intent detection to be general to dialogue meta-train datasets lacking widely-varying domains, which focuses on the scope of episodic intent classes to construct meta-task dynamically.

1527, TITLE: Conversational Query Rewriting with Self-Supervised Learning

<https://doi.org/10.1109/ICASSP39728.2021.9413557>

AUTHORS: H. Liu, M. Chen, Y. Wu, X. He and B. Zhou

HIGHLIGHT: To tackle these issues, we first propose to construct a large-scale CQR dataset automatically via self-supervised learning, which does not need human annotation. Then we introduce a novel CQR model Teresa based on Transformer, which is enhanced by self-attentive keywords detection and intent consistency constraint.

1528, TITLE: Handling Class Imbalance in Low-Resource Dialogue Systems by Combining Few-Shot Classification and Interpolation

<https://doi.org/10.1109/ICASSP39728.2021.9413405>

AUTHORS: V. Sunder and E. Fosler-Lussier

HIGHLIGHT: We present a new end-to-end pairwise learning framework that is designed specifically to tackle this phenomenon by inducing a few-shot classification capability in the utterance representations and augmenting data through an interpolation of utterance representations.

1529, TITLE: Improving Cross-Domain Slot Filling with Common Syntactic Structure

<https://doi.org/10.1109/ICASSP39728.2021.9414625>

AUTHORS: L. Liu, X. Lin, P. Zhang and B. Wang

HIGHLIGHT: In this paper, we attempt to solve this task by exploiting the syntactic structures of user utterances, because these syntactic structures are actually accessible and can be shared between utterances from different domains.

1530, TITLE: Joint Intent Detection and Slot Filling Based on Continual Learning Model

<https://doi.org/10.1109/ICASSP39728.2021.9413360>



AUTHORS: Y. Hui, J. Wang, N. Cheng, F. Yu, T. Wu and J. Xiao  
HIGHLIGHT: In this paper, a Continual Learning Interrelated Model (CLIM) is proposed to consider semantic information with different characteristics and balance the accuracy between intent detection and slot filling effectively.

1531, TITLE: Knowledge-Based Chat Detection with False Mention Discrimination  
<https://doi.org/10.1109/ICASSP39728.2021.9414073>  
AUTHORS: W. Liu, P. Huang, D. Liang and Z. Zhou  
HIGHLIGHT: To deal with this issue, this paper proposes a new model for knowledge-based chat detection with false mention discrimination (FMD-KChat).

1532, TITLE: Replacing Human Audio with Synthetic Audio for on-Device Unspoken Punctuation Prediction  
<https://doi.org/10.1109/ICASSP39728.2021.9413432>  
AUTHORS: D. Soboleva et al.  
HIGHLIGHT: We present a novel multi-modal unspoken punctuation prediction system for the English language which combines acoustic and text features.

1533, TITLE: Adversarial Generative Distance-Based Classifier for Robust Out-of-Domain Detection  
<https://doi.org/10.1109/ICASSP39728.2021.9413908>  
AUTHORS: Z. Zeng et al.  
HIGHLIGHT: In this paper, we propose an efficient adversarial attack mechanism to augment hard OOD samples and design a novel generative distance-based classifier to detect OOD samples instead of a traditional threshold-based discriminator classifier.

1534, TITLE: GAN-Based Out-of-Domain Detection Using Both In-Domain and Out-of-Domain Samples  
<https://doi.org/10.1109/ICASSP39728.2021.9414246>  
AUTHORS: C. Liang, P. Huang, W. Lai and Z. Ruan  
HIGHLIGHT: In the situation where both in-domain (ID) and OOD samples are available, our goal is to take advantage of OOD samples under the GAN-based framework for OOD detection.

1535, TITLE: Progressive Dialogue State Tracking for Multi-Domain Dialogue Systems  
<https://doi.org/10.1109/ICASSP39728.2021.9414610>  
AUTHORS: J. Wang, M. Liu and X. Quan  
HIGHLIGHT: To model the two observations, we propose to divide the task into two successive procedures: progressive domain-slot tracking and shrunk value prediction.

1536, TITLE: Multi-Step Spoken Language Understanding System Based on Adversarial Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9414514>  
AUTHORS: Y. Wang, Y. Shen and H. Jin  
HIGHLIGHT: In this paper, we introduce a novel multi-step spoken language understanding system based on adversarial learning that can leverage the multiround user's feedback to update slot values.

1537, TITLE: Multi-Entity Collaborative Relation Extraction  
<https://doi.org/10.1109/ICASSP39728.2021.9413673>  
AUTHORS: H. Liu, Z. Li, D. Sheng, H. -T. Zheng and Y. Shen  
HIGHLIGHT: Instead of designing specific models for single relationship extraction tasks, this paper aims to propose a general framework to extract multiple relations among multiple entities in unstructured text by taking advantage of existing models.

1538, TITLE: Multi-Granularity Heterogeneous Graph for Document-Level Relation Extraction  
<https://doi.org/10.1109/ICASSP39728.2021.9414755>  
AUTHORS: H. Tang, Y. Cao, Z. Zhang, R. Jia, F. Fang and S. Wang  
HIGHLIGHT: In this paper, we propose a novel Multi-granularity Heterogeneous Graph (MHG) to tackle this challenge.

1539, TITLE: Improving Event Detection by Exploiting Label Hierarchy  
<https://doi.org/10.1109/ICASSP39728.2021.9415002>  
AUTHORS: X. Xi et al.  
HIGHLIGHT: To fully utilize such information to improve the detection of fine-grained event types, we propose a three-layer label hierarchy and introduce the detection of two coarser-grained types as auxiliary classification tasks.

1540, TITLE: Improving NER in Social Media via Entity Type-Compatible Unknown Word Substitution  
<https://doi.org/10.1109/ICASSP39728.2021.9414304>  
AUTHORS: J. Xie, K. Zhang, L. Sun, Y. Su and C. Xu

**HIGHLIGHT:** To alleviate the noisy expression in social media data, we present a novel word substitution strategy based on constructing an entity type-compatible (ETC) semantic space.

1541, **TITLE:** More: A Metric Learning Based Framework for Open-Domain Relation Extraction

<https://doi.org/10.1109/ICASSP39728.2021.9413437>

**AUTHORS:** Y. Wang, R. Lou, K. Zhang, M. Y. Chen and Y. Yang

**HIGHLIGHT:** To address these problems, in this work, we propose a novel learning framework named MORE (Metric learning-based Open Relation Extraction).

1542, **TITLE:** ?You Should Probably Read This?: Hedge Detection in Text

<https://doi.org/10.1109/ICASSP39728.2021.9414095>

**AUTHORS:** D. Katerenchuk and R. Levitan

**HIGHLIGHT:** In this work, we apply a joint model that leverages words and part-of-speech tags to improve hedge detection in text and achieve a new top score on the CoNLL-2010 Wikipedia corpus.

1543, **TITLE:** Enhancing Model Robustness by Incorporating Adversarial Knowledge into Semantic Representation

<https://doi.org/10.1109/ICASSP39728.2021.9414793>

**AUTHORS:** J. Li, T. Du, X. Liu, R. Zhang, H. Xue and S. Ji

**HIGHLIGHT:** Therefore, we propose AdvGraph, a novel defense which enhances the robustness of Chinese-based NLP models by incorporating adversarial knowledge into the semantic representation of the input.

1544, **TITLE:** Elbert: Fast Albert with Confidence-Window Based Early Exit

<https://doi.org/10.1109/ICASSP39728.2021.9414572>

**AUTHORS:** K. Xie, S. Lu, M. Wang and Z. Wang

**HIGHLIGHT:** In this work, we propose the ELBERT, which significantly improves the average inference speed compared to ALBERT due to the proposed confidence-window based early exit mechanism, without introducing additional parameters or extra training overhead.

1545, **TITLE:** Dualformer: A Unified Bidirectional Sequence-to-Sequence Learning

<https://doi.org/10.1109/ICASSP39728.2021.9413402>

**AUTHORS:** J. -T. Chien and W. -H. Chang

**HIGHLIGHT:** This paper presents a new dual domain mapping based on a unified bidirectional sequence-to-sequence (seq2seq) learning.

1546, **TITLE:** Task Aware Multi-Task Learning for Speech to Text Tasks

<https://doi.org/10.1109/ICASSP39728.2021.9414703>

**AUTHORS:** S. Indurthi et al.

**HIGHLIGHT:** We propose a task modulation network which allows the model to learn task specific features, while learning the shared features simultaneously.

1547, **TITLE:** Label-Aware Text Representation for Multi-Label Text Classification

<https://doi.org/10.1109/ICASSP39728.2021.9413921>

**AUTHORS:** H. Guo, X. Li, L. Zhang, J. Liu and W. Chen

**HIGHLIGHT:** In this paper, we propose a label-aware network to obtain both the label correlation and text representation.

1548, **TITLE:** Mixup Regularized Adversarial Networks for Multi-Domain Text Classification

<https://doi.org/10.1109/ICASSP39728.2021.9413441>

**AUTHORS:** Y. Wu, D. Inkpen and A. El-Roby

**HIGHLIGHT:** In this paper, we propose mixup regularized adversarial networks (MRANs) to address these two issues.

1549, **TITLE:** Mispronunciation Detection in Non-Native (L2) English with Uncertainty Modeling

<https://doi.org/10.1109/ICASSP39728.2021.9413953>

**AUTHORS:** D. Korzekwa, J. Lorenzo-Trueba, S. Zaporowski, S. Calamaro, T. Drugman and B. Kostek

**HIGHLIGHT:** We propose a novel approach to overcome this problem based on two principles: a) taking into account uncertainty in the automatic phoneme recognition step, b) accounting for the fact that there may be multiple valid pronunciations.

1550, **TITLE:** Attention-Based Multi-Encoder Automatic Pronunciation Assessment

<https://doi.org/10.1109/ICASSP39728.2021.9414451>

**AUTHORS:** B. Lin and L. Wang

**HIGHLIGHT:** In this paper we propose an end-to-end (E2E) pronunciation scoring network based on attention mechanism and multi-encoder consisting of audio and text encoders.

1551, **TITLE:** Improving Pronunciation Assessment Via Ordinal Regression with Anchored Reference Samples  
<https://doi.org/10.1109/ICASSP39728.2021.9413659>

**AUTHORS:** B. Su, S. Mao, F. Soong, Y. Xia, J. Tien and Z. Wu

**HIGHLIGHT:** In this paper, we propose two new statistical features, average GOP (aGOP) and confusion GOP (cGOP) and use them to train a binary classifier in Ordinal Regression with Anchored Reference Samples (ORARS).

1552, **TITLE:** Analysing Bias in Spoken Language Assessment Using Concept Activation Vectors  
<https://doi.org/10.1109/ICASSP39728.2021.9413988>

**AUTHORS:** X. Wei, M. J. F. Gales and K. M. Knill

**HIGHLIGHT:** This work applies CAVs to assess bias in a spoken language assessment (SLA) system, a regression task.

1553, **TITLE:** Senone-Aware Adversarial Multi-Task Training for Unsupervised Child to Adult Speech Adaptation  
<https://doi.org/10.1109/ICASSP39728.2021.9413738>

**AUTHORS:** R. Duan and N. F. Chen

**HIGHLIGHT:** In this work, we propose a feature adaptation approach by exploiting adversarial multi-task training to minimize acoustic mismatch at the senone (tied triphone states) level between adult and child speech and leverage large amounts of transcribed adult speech.

1554, **TITLE:** Classifying Speech Intelligibility Levels of Children in Two Continuous Speech Styles  
<https://doi.org/10.1109/ICASSP39728.2021.9414986>

**AUTHORS:** Y. -S. Lin and S. -C. Tseng

**HIGHLIGHT:** This paper attempts to construct automatic systems that help detect children with severe speech problems at an early stage.

1555, **TITLE:** Recent Advances in Arabic Syntactic Diacritics Restoration  
<https://doi.org/10.1109/ICASSP39728.2021.9414500>

**AUTHORS:** Y. Hifny

**HIGHLIGHT:** In this paper, we implement a knowledge distillation technique where an ensemble of teachers/taggers is used to train a single student tagger.

1556, **TITLE:** Making Punctuation Restoration Robust and Fast with Multi-Task Learning and Knowledge Distillation  
<https://doi.org/10.1109/ICASSP39728.2021.9414518>

**AUTHORS:** M. Hentschel, E. Tsunoo and T. Okuda

**HIGHLIGHT:** To address the former, we use a multi-task learning framework with ELECTRA, a recently proposed improvement on BERT, that has a generator-discriminator structure.

1557, **TITLE:** Variational Dialogue Generation with Normalizing Flows  
<https://doi.org/10.1109/ICASSP39728.2021.9414586>

**AUTHORS:** T. -C. Luo and J. -T. Chien

**HIGHLIGHT:** This paper presents the dialogue flow VAE (DF-VAE) for variational dialogue generation.

1558, **TITLE:** NN-KOG2P: A Novel Grapheme-to-Phoneme Model for Korean Language  
<https://doi.org/10.1109/ICASSP39728.2021.9414653>

**AUTHORS:** H. -Y. Kim, J. -H. Kim and J. -M. Kim

**HIGHLIGHT:** In this paper, we propose a novel Korean G2P model architecture, reflecting the characteristics of Korean pronunciation, called neural network-based Korean G2P (NN-KoG2P).

1559, **TITLE:** Joint Alignment Learning-Attention Based Model for Grapheme-to-Phoneme Conversion  
<https://doi.org/10.1109/ICASSP39728.2021.9413679>

**AUTHORS:** Y. Wang, F. Bao, H. Zhang and G. Gao

**HIGHLIGHT:** In this paper, we present a novel approach to optimize the G2P conversion model directly alignment grapheme-phoneme sequence by using alignment learning (AL) as the loss function.

1560, **TITLE:** Knowledge Distillation for Improved Accuracy in Spoken Question Answering  
<https://doi.org/10.1109/ICASSP39728.2021.9414999>

**AUTHORS:** C. You, N. Chen and Y. Zou

**HIGHLIGHT:** However, the recent work shows that ASR systems generate highly noisy transcripts, which critically limit the capability of machine comprehension on the SQA task. To address the issue, we present a novel distillation framework.

1561, **TITLE:** Coarse-To-Careful: Seeking Semantic-Related Knowledge for Open-Domain Commonsense Question Answering

<https://doi.org/10.1109/ICASSP39728.2021.9413878>

**AUTHORS:** L. Xing, Y. Hu, J. Yu, Y. Xie and W. Peng

**HIGHLIGHT:** Towards the issue of introducing related knowledge, we propose a semantic-driven knowledge-aware QA framework, which controls the knowledge injection in a coarse-to-careful fashion.

1562, **TITLE:** Language Model is all You Need: Natural Language Understanding as Question Answering

<https://doi.org/10.1109/ICASSP39728.2021.9413810>

**AUTHORS:** M. Namazifar, A. Papangelis, G. Tur and D. Hakkani-T?r

**HIGHLIGHT:** In this work we study the use of a certain family of transfer learning, where the target domain is mapped to the source domain.

1563, **TITLE:** Integrating Subgraph-Aware Relation and Direction Reasoning for Question Answering

<https://doi.org/10.1109/ICASSP39728.2021.9413620>

**AUTHORS:** X. Wang et al.

**HIGHLIGHT:** To address these challenges, we propose a novel neural model, Relation-updated Direction-guided Answer Selector (RDAS), which converts relations in each subgraph to additional nodes to learn structure information.

1564, **TITLE:** Role Aware Multi-Party Dialogue Question Answering

<https://doi.org/10.1109/ICASSP39728.2021.9414980>

**AUTHORS:** J. -H. Hsu, P. -W. Shen, H. -T. Su, C. -H. Chang, J. -F. Yeh and W. H. Hsu

**HIGHLIGHT:** In a novel aspect, this paper proposes the Role Aware Multi-Party Network (RAMPNet), a model utilizing the information of speaker and role to present "who is speaking" and "who is mentioned", making role awareness an available message for our model.

1565, **TITLE:** MCR-NET: A Multi-Step Co-Interactive Relation Network for Unanswerable Questions on Machine Reading Comprehension

<https://doi.org/10.1109/ICASSP39728.2021.9413435>

**AUTHORS:** W. Peng et al.

**HIGHLIGHT:** To tackle this problem, we propose a Multi-Step Co-Interactive Relation Network (MCR-Net) to explicitly model the mutual interaction and locate key clues from coarse to fine by introducing a co-interactive relation module.

1566, **TITLE:** Hierarchical Speaker-Aware Sequence-to-Sequence Model for Dialogue Summarization

<https://doi.org/10.1109/ICASSP39728.2021.9414547>

**AUTHORS:** Y. Lei, Y. Yan, Z. Zeng, K. He, X. Zhang and W. Xu?

**HIGHLIGHT:** In this paper, we propose a hierarchical transformer-based model for dialogue summarization.

1567, **TITLE:** A Large-Scale Chinese Long-Text Extractive Summarization Corpus

<https://doi.org/10.1109/ICASSP39728.2021.9414946>

**AUTHORS:** K. Chen, G. Fu, Q. Chen and B. Hu

**HIGHLIGHT:** In this paper, we publish a large-scale Chinese Long-text Extractive Summarization corpus named CLES.

1568, **TITLE:** Adaptive Bi-Directional Attention: Exploring Multi-Granularity Representations for Machine Reading Comprehension

<https://doi.org/10.1109/ICASSP39728.2021.9414067>

**AUTHORS:** N. Chen, F. Liu, C. You, P. Zhou and Y. Zou

**HIGHLIGHT:** To this end, we propose a novel approach called Adaptive Bidirectional Attention, which adaptively exploits the source representations of different levels to the predictor.

1569, **TITLE:** Graph Attention and Interaction Network With Multi-Task Learning for Fact Verification

<https://doi.org/10.1109/ICASSP39728.2021.9414889>

**AUTHORS:** R. Yang, R. Wang and Z. -H. Ling

**HIGHLIGHT:** In this paper, we propose a graph attention and interaction network (GAIN) for claim verification.

1570, **TITLE:** Enhancing Deep Paraphrase Identification via Leveraging Word Alignment Information

<https://doi.org/10.1109/ICASSP39728.2021.9414944>

- AUTHORS: B. Li, T. Liu, B. Wang and L. Wang  
HIGHLIGHT: Apart from directly encoding WAI into fixed-size embeddings, we propose a novel auxiliary task so that the baselines can be pre-trained using a large amount of unlabeled in-domain data.
- 1571, TITLE: An End-To-End Actor-Critic-Based Neural Coreference Resolution System  
<https://doi.org/10.1109/ICASSP39728.2021.9413579>  
AUTHORS: Y. Wang, Y. Shen and H. Jin  
HIGHLIGHT: In this paper, we propose an actor-critic-based neural coreference resolution system, which can achieve both mention detection and mention clustering by leveraging an actor-critic deep reinforcement learning technique and a joint training algorithm.
- 1572, TITLE: Reduced-Complexity Modular Polynomial Multiplication for R-LWE Cryptosystems  
<https://doi.org/10.1109/ICASSP39728.2021.9414005>  
AUTHORS: X. Zhang and K. K. Parhi  
HIGHLIGHT: In this paper, a new method is proposed to integrate the modular reduction into the Karatsuba polynomial multiplication.
- 1573, TITLE: Seizure Detection Using Power Spectral Density via Hyperdimensional Computing  
<https://doi.org/10.1109/ICASSP39728.2021.9414083>  
AUTHORS: L. Ge and K. K. Parhi  
HIGHLIGHT: This paper explores seizure detection from electroencephalogram (EEG) from subjects with epilepsy using HD computing based on power spectral density (PSD) features.
- 1574, TITLE: FPGA Hardware Design for Plenoptic 3D Image Processing Algorithm Targeting a Mobile Application  
<https://doi.org/10.1109/ICASSP39728.2021.9414690>  
AUTHORS: F. Bhatti and T. Greiner  
HIGHLIGHT: The idea presented in this paper is to use the FPGA based hardware design to improve the performance of a 3D depth estimation algorithm by utilizing the advantage of concurrent execution.
- 1575, TITLE: SLAP: a Split Latency Adaptive VLIW Pipeline Architecture Which Enables on-The-Fly Variable SIMD Vector-Length  
<https://doi.org/10.1109/ICASSP39728.2021.9413371>  
AUTHORS: A. Shrivastava, A. Gatherer, T. Sun, S. Wokhlu and A. Chandra  
HIGHLIGHT: We describe the Split Latency Adaptive Pipeline (SLAP) VLIW architecture, a cache performance improvement technology that requires zero change to object code, while removing smart DMAs and their overhead.
- 1576, TITLE: Unsupervised Clustering of Time Series Signals Using Neuromorphic Energy-Efficient Temporal Neural Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414882>  
AUTHORS: S. Chaudhari, H. Nair, J. M. F. Moura and J. Paul Shen  
HIGHLIGHT: In this work, we propose a neuromorphic approach to unsupervised time series clustering based on Temporal Neural Networks that is capable of ultra low-power, continuous online learning.
- 1577, TITLE: Angle-of-Arrival (AoA) Factorization in Multipath Channels  
<https://doi.org/10.1109/ICASSP39728.2021.9414026>  
AUTHORS: Y. -L. Wei and R. R. Choudhury  
HIGHLIGHT: This paper considers the problem of estimating K angle of arrivals (AoA) using an array of  $M > K$  microphones.
- 1578, TITLE: Scaled Fast Nested Key Equation Solver for Generalized Integrated Interleaved BCH Decoders  
<https://doi.org/10.1109/ICASSP39728.2021.9414846>  
AUTHORS: Z. Xie and X. Zhang  
HIGHLIGHT: This paper proposes novel reformulations of the nested BCH KES to enable scalar pre-computation.
- 1579, TITLE: Joint Optimization for Full-Duplex Cellular Communications Via Intelligent Reflecting Surface  
<https://doi.org/10.1109/ICASSP39728.2021.9413615>  
AUTHORS: Z. Peng, C. Pan, Z. Zhang, X. Chen, L. Li and A. L. Swindlehurst  
HIGHLIGHT: Specifically, we propose a low-complexity minorization-maximization (MM) algorithm for solving the subproblems of designing the precoding matrix and the reflection coefficients, respectively.
- 1580, TITLE: A Color Doppler Processing Engine with an Adaptive Clutter Filter for Portable Ultrasound Imaging Devices

<https://doi.org/10.1109/ICASSP39728.2021.9414344>

AUTHORS: Y. -L. Lo and C. -H. Yang

HIGHLIGHT: This paper presents an optimized color Doppler processing engine for portable ultrasound devices.

1581, TITLE: Convolutional Neural Network-Aided Bit-Flipping for Belief Propagation Decoding of Polar Codes

<https://doi.org/10.1109/ICASSP39728.2021.9413808>

AUTHORS: C. -F. Teng, A. K. -S. Ho, C. -H. D. Wu, S. -S. Wong and A. -Y. A. Wu

HIGHLIGHT: In this work, we propose a convolutional neural network-aided bit-flipping (CNN-BF) mechanism to further enhance BP decoding.

1582, TITLE: Taming Voting Algorithms on Gpus for an Efficient Connected Component Analysis Algorithm

<https://doi.org/10.1109/ICASSP39728.2021.9413653>

AUTHORS: F. Lemaitre, A. Hennequin and L. Lacassagne

HIGHLIGHT: This paper explores multiple ways to reduce those conflicts for voting algorithms and especially for Connected Component Analysis.

1583, TITLE: Positnn: Training Deep Neural Networks with Mixed Low-Precision Posit

<https://doi.org/10.1109/ICASSP39728.2021.9413919>

AUTHORS: G. Raposo, P. Tom's and N. Roma

HIGHLIGHT: The presented research aims to evaluate the feasibility to train deep convolutional neural networks using posits.

1584, TITLE: Bluetooth Low Energy and CNN-Based Angle of Arrival Localization in Presence of Rayleigh Fading

<https://doi.org/10.1109/ICASSP39728.2021.9413455>

AUTHORS: Z. HajiAkhondi-Meybodi, M. Salimibeni, A. Mohammadi and K. N. Plataniotis

HIGHLIGHT: The paper proposes an efficient Convolutional Neural Network (CNN)-based indoor localization framework to tackle these issues specific to BLE-based settings.

1585, TITLE: Robust Device-Free Proximity Detection Using Wifi

<https://doi.org/10.1109/ICASSP39728.2021.9414707>

AUTHORS: Y. Hu, M. Z. Ozturk, F. Zhang, B. Wang and K. J. Ray Liu

HIGHLIGHT: In this paper, we propose two robust and responsive features in the frequency dimension, which are sensitive to the distance of motion, and establish the connection between the underlying radio propagation properties and the features.

1586, TITLE: Online Dynamic Window (ODW) Assisted 2-Stage LSTM Indoor Localization for Smart Phones

<https://doi.org/10.1109/ICASSP39728.2021.9414174>

AUTHORS: M. Atashi and A. Mohammadi

HIGHLIGHT: In this regard, the paper takes one step forward to transfer offline IMU-based models to online positioning frameworks.

1587, TITLE: Optimal TOA Localization for Moving Sensor in Asymmetric Network

<https://doi.org/10.1109/ICASSP39728.2021.9414962>

AUTHORS: S. Zhao, X. -P. Zhang, X. Cui and M. Lu

HIGHLIGHT: We develop an optimal localization method based-on maximum likelihood (ML) estimator, namely ML-LOC, utilizing information on the SN velocity and clock drift, to determine the position of a moving SN.

1588, TITLE: Low Complexity SLM for OFDMA System with Implicit Side Information

<https://doi.org/10.1109/ICASSP39728.2021.9414599>

AUTHORS: S. Hu, M. Yang, K. Kang and H. Qian

HIGHLIGHT: In this paper, we propose a novel SLM algorithm for the OFDMA system that requires no side information transmission.

1589, TITLE: Reduced-Complexity Channel Estimation by Hierarchical Interpolation Exploiting Sparsity for Massive MIMO Systems with Uniform Rectangular Array

<https://doi.org/10.1109/ICASSP39728.2021.9414289>

AUTHORS: C. -S. Wang and P. -Y. Tsai

HIGHLIGHT: To save the complexity, we propose hierarchical channel interpolation algorithm by exploiting the channel sparsity in the millimeter wave frequency band.

1590, TITLE: Traffic Speed Forecasting Via Spatio-Temporal Attentive Graph Isomorphism Network

<https://doi.org/10.1109/ICASSP39728.2021.9414596>

- AUTHORS: Q. Yang, T. Zhong and F. Zhou  
HIGHLIGHT: This work proposes an end-to-end framework to capture spatial dependencies through graph isomorphism network, while explicitly taking network topologic similarities into account and leveraging symmetric traffic for learning the traffic conditions.
- 1591, TITLE: Inferring High-Resolutional Urban Flow With Internet Of Mobile Things  
<https://doi.org/10.1109/ICASSP39728.2021.9414134>  
AUTHORS: F. Zhou, X. Jing, L. Li and T. Zhong  
HIGHLIGHT: This work introduces a new method for inferring fine-grained urban flow with the internet of mobile things such as taxis and bikes.
- 1592, TITLE: Transfer Learning for Input Estimation of Vehicle Systems  
<https://doi.org/10.1109/ICASSP39728.2021.9413671>  
AUTHORS: L. M. Cronin, S. S. Eshkevari, D. Sen and S. N. Pakzad  
HIGHLIGHT: This study proposes a learning-based method with domain adaptability for input estimation of vehicle suspension systems.
- 1593, TITLE: Identification of Deep Breath While Moving Forward Based on Multiple Body Regions and Graph Signal Analysis  
<https://doi.org/10.1109/ICASSP39728.2021.9413546>  
AUTHORS: Y. Wang et al.  
HIGHLIGHT: This paper presents an unobtrusive solution that can automatically identify deep breath when a person is walking past the global depth camera.
- 1594, TITLE: Multi-Object Tracking Using Poisson Multi-Bernoulli Mixture Filtering For Autonomous Vehicles  
<https://doi.org/10.1109/ICASSP39728.2021.9415072>  
AUTHORS: S. Pang and H. Radha  
HIGHLIGHT: In this work, we developed an RFS-based MOT framework for 3D LiDAR data.
- 1595, TITLE: Adaptive RF Fingerprint Decomposition in Micro UAV Detection based on Machine Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9414985>  
AUTHORS: C. Xu, F. He, B. Chen, Y. Jiang and H. Song  
HIGHLIGHT: This paper proposes a method using empirical mode decomposition (EMD) and ensemble empirical mode decomposition (EEMD) on extracting the communication channel characteristics of intruding UAVs.
- 1596, TITLE: Depression Detection by Analysing Eye Movements on Emotional Images  
<https://doi.org/10.1109/ICASSP39728.2021.9414663>  
AUTHORS: R. Shen, Q. Zhan, Y. Wang and H. Ma  
HIGHLIGHT: To achieve an objective and efficient depression detection system, we propose a cognitive psychology experimental paradigm based on the attentional bias theory and eye movements in this paper.
- 1597, TITLE: Weakly Supervised Patch Label Inference Network with Image Pyramid for Pavement Diseases Recognition in the Wild  
<https://doi.org/10.1109/ICASSP39728.2021.9413517>  
AUTHORS: G. Huang, S. Huang, L. Huangfu and D. Yang  
HIGHLIGHT: In this paper, we present an end-to-end deep learning approach named Weakly Super-vised Patch Label Inference Network with Image Pyramid (WSPLIN-IP) for recognizing various types of pavement diseases that are not just limited to the specific ones, such as crack and pothole.
- 1598, TITLE: A Wireless Reference Active Noise Control Headphone Using Coherence Based Selection Technique  
<https://doi.org/10.1109/ICASSP39728.2021.9414683>  
AUTHORS: X. Shen, D. Shi and W. -S. Gan  
HIGHLIGHT: Hence, we adopt a wireless reference microphone to pick up the reference signals around the noise sources.
- 1599, TITLE: How to Use Time Information Effectively? Combining with Time Shift Module for Lipreading  
<https://doi.org/10.1109/ICASSP39728.2021.9414659>  
AUTHORS: M. Hao, M. Mamut, N. Yadikar, A. Aysa and K. Ubul  
HIGHLIGHT: The proposed method verified on two challenging word-level lipreading datasets LRW and LRW-1000 and achieved new state-of-the-art performance.



- 1600, TITLE: Exploring the application of synthetic audio in training keyword spotters  
<https://doi.org/10.1109/ICASSP39728.2021.9413448>  
AUTHORS: A. Werchniak et al.  
HIGHLIGHT: This paper details some initial exploration into the application of Text-To-Speech (TTS) audio as a "helper" tool for training keyword spotters in these low-resource scenarios.
- 1601, TITLE: Graph Enhanced Query Rewriting for Spoken Language Understanding System  
<https://doi.org/10.1109/ICASSP39728.2021.9413840>  
AUTHORS: S. Yuan, S. Gupta, X. Fan, D. Liu, Y. Liu and C. Guo  
HIGHLIGHT: In this work, we construct a user interaction graph from their queries using data mined from a Markov Chain Model [1], and introduce a self-supervised pre-training process for learning query embeddings by leveraging the recent developments in Graph Representation Learning (GRL).
- 1602, TITLE: Deep Neural Network Based Cough Detection Using Bed-Mounted Accelerometer Measurements  
<https://doi.org/10.1109/ICASSP39728.2021.9414744>  
AUTHORS: M. Pahar, I. Miranda, A. Diacon and T. Niesler  
HIGHLIGHT: We have performed cough detection based on measurements from an accelerometer attached to the patient's bed.
- 1603, TITLE: Radio Frequency Based Heart Rate Variability Monitoring  
<https://doi.org/10.1109/ICASSP39728.2021.9413465>  
AUTHORS: F. Wang, X. Zeng, C. Wu, B. Wang and K. J. Ray Liu  
HIGHLIGHT: In this paper, we present mmHRV, a contact-free HRV monitoring system using commercial millimeter-wave (mmWave) radio.
- 1604, TITLE: Discrete Cosine Transform Based Causal Convolutional Neural Network for Drift Compensation in Chemical Sensors  
<https://doi.org/10.1109/ICASSP39728.2021.9414512>  
AUTHORS: D. Badawi, A. Agambayev, S. Ozev and A. Enis ?etin  
HIGHLIGHT: In this paper, we develop a causal convolutional neural network (CNN) with a Discrete Cosine Transform (DCT) layer to estimate the drift signal.
- 1605, TITLE: Wifi-Based Device-Free Gesture Recognition Through-the-Wall  
<https://doi.org/10.1109/ICASSP39728.2021.9414894>  
AUTHORS: S. D. Regani, B. Wang and K. J. Ray Liu  
HIGHLIGHT: In this work, we propose such a gesture recognition system that can recover information about the actual trajectory of the hand movement allowing an expandable set of gestures.
- 1606, TITLE: Sound Recovery From Radio Signals  
<https://doi.org/10.1109/ICASSP39728.2021.9413508>  
AUTHORS: M. Z. Ozturk, C. Wu, B. Wang and K. J. Ray Liu  
HIGHLIGHT: In this work, we model the vibration on object surfaces due to sound for mmWave devices.
- 1607, TITLE: Fully-Neural Approach to Vehicle Weighing and Strain Prediction on Bridges Using Wireless Accelerometers  
<https://doi.org/10.1109/ICASSP39728.2021.9414433>  
AUTHORS: T. Kawakatsu, K. Aihara, A. Takasu, J. Adachi, H. Wang and T. Nagayama  
HIGHLIGHT: In this paper, we propose a new BWIM approach based on a deep neural network using accelerometers, which are easier to install than strain sensors, thus helping the advancement of low-cost BWIM systems.
- 1608, TITLE: End To End Learning For Convolutional Multi-Channel Wiener Filtering  
<https://doi.org/10.1109/ICASSP39728.2021.9414907>  
AUTHORS: M. Togami  
HIGHLIGHT: In this paper, we propose a dereverberation and speech source separation method based on deep neural network (DNN).
- 1609, TITLE: Makf-Sr: Multi-Agent Adaptive Kalman Filtering-Based Successor Representations  
<https://doi.org/10.1109/ICASSP39728.2021.9414597>  
AUTHORS: M. Salimibeni, P. Malekzadeh, A. Mohammadi, P. Spachos and K. N. Plataniotis  
HIGHLIGHT: The proposed framework can adapt quickly to the changes in a multi-agent environment faster than the MF methods and with a lower computational cost compared to MB algorithms.

- 1610, TITLE: Variation-Stable Fusion for PPG-Based Biometric System  
<https://doi.org/10.1109/ICASSP39728.2021.9413906>  
AUTHORS: D. Y. Hwang, B. Taha and D. Hatzinakos  
HIGHLIGHT: This paper investigates the employment of photoplethysmography (PPG) for user authentication systems.
- 1611, TITLE: Improving Stability of Adversarial Li-ion Cell Usage Data Generation using Generative Latent Space Modelling  
<https://doi.org/10.1109/ICASSP39728.2021.9413892>  
AUTHORS: S. Chatteraj, S. Pratiher, S. Pratiher and H. Konik  
HIGHLIGHT: In this treatise, some robust loss-functions for the TimeGAN architecture are explored for generating realistic Li-ion CUD.
- 1612, TITLE: SQWA: Stochastic Quantized Weight Averaging For Improving The Generalization Capability Of Low-Precision Deep Neural Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9413623>  
AUTHORS: S. Shin, Y. Boo and W. Sung  
HIGHLIGHT: We present a new quantized neural network optimization approach, stochastic quantized weight averaging (SQWA), to design low-precision DNNs with good generalization capability using model averaging.
- 1613, TITLE: A Quantitative Analysis Of The Robustness Of Neural Networks For Tabular Data  
<https://doi.org/10.1109/ICASSP39728.2021.9413858>  
AUTHORS: K. Gupta, B. Pesquet-Popescu, F. Kaakai and J. -C. Pesquet  
HIGHLIGHT: This paper presents a quantitative approach to demonstrate the robustness of neural networks for tabular data.
- 1614, TITLE: Spatial Equalization Before Reception: Reconfigurable Intelligent Surfaces for Multi-Path Mitigation  
<https://doi.org/10.1109/ICASSP39728.2021.9414612>  
AUTHORS: H. Zhang, L. Song, Z. Han and H. V. Poor  
HIGHLIGHT: In this paper, we propose to use an RIS as a spatial equalizer to address the well-known multi-path fading phenomenon.
- 1615, TITLE: Interference Analysis in Reconfigurable Intelligent Surface-Assisted Multiple-Input Multiple-Output Systems  
<https://doi.org/10.1109/ICASSP39728.2021.9414684>  
AUTHORS: J. Liu, X. Qian and M. Di Renzo  
HIGHLIGHT: In this paper, we consider a multiple-input multiple-output network where each base station serves a user equipment with the aid of an RIS equipped with N reconfigurable elements.
- 1616, TITLE: Codebook Design for Dual-Polarized Ultra-Massive MIMO Communications at Millimeter Wave and Terahertz Bands  
<https://doi.org/10.1109/ICASSP39728.2021.9413660>  
AUTHORS: S. Nie and I. F. Akyildiz  
HIGHLIGHT: In this work, we analyze the dual-polarized intelligent surface channel and present an efficient precoding solution based on the array-of-subarray architecture that aims to maximize the spectral efficiency.
- 1617, TITLE: Performance Analysis of Spatial and Frequency Domain Index-Modulated Reconfigurable Intelligent Metasurfaces  
<https://doi.org/10.1109/ICASSP39728.2021.9413387>  
AUTHORS: J. A. Hodge, K. V. Mishra, B. M. Sadler and A. I. Zaghoul  
HIGHLIGHT: In this paper, we propose novel electromagnetics-compliant designs of reconfigurable intelligent surface (RIS) apertures for realizing IM in 6G transceivers.
- 1618, TITLE: Meta-Learning for 6G Communication Networks with Reconfigurable Intelligent Surfaces  
<https://doi.org/10.1109/ICASSP39728.2021.9413598>  
AUTHORS: M. Jung and W. Saad  
HIGHLIGHT: In this paper, a practical channel acquisition and passive beamforming technique is proposed using a limited number of pilot symbols in an RIS-assisted cellular network.
- 1619, TITLE: Model-Inspired Deep Learning for Light-Field Microscopy with Application to Neuron Localization  
<https://doi.org/10.1109/ICASSP39728.2021.9414236>  
AUTHORS: P. Song, H. V. Jadan, C. L. Howe, P. Quicke, A. J. Foust and P. Luigi Dragotti

**HIGHLIGHT:** In this work, we propose a model-inspired deep learning approach to perform fast and robust 3D localization of sources using light-field microscopy images.

1620, **TITLE:** Time-Varying Graph Signal Inpainting Via Unrolling Networks

<https://doi.org/10.1109/ICASSP39728.2021.9413406>

**AUTHORS:** S. Chen and Y. C. Eldar

**HIGHLIGHT:** We propose an interpretable graph neural network based on algorithm unrolling to reconstruct a time-varying graph signal from partial measurements.

1621, **TITLE:** Deep Learning for Linear Inverse Problems Using the Plug-and-Play Priors Framework

<https://doi.org/10.1109/ICASSP39728.2021.9413947>

**AUTHORS:** W. Chen, D. Wipf and M. Rodrigues

**HIGHLIGHT:** In this overview paper, we present the combination of the DL and the Plug-and-Play priors (PPP) framework, showcasing how it allows solving various inverse problems by leveraging the impressive capabilities of existing DL based denoising algorithms.

1622, **TITLE:** A Plug-and-Play Deep Image Prior

<https://doi.org/10.1109/ICASSP39728.2021.9414879>

**AUTHORS:** Z. Sun, F. Latorre, T. Sanchez and V. Cevher

**HIGHLIGHT:** To mitigate this effect, this work incorporates a plug-and-play prior scheme which can accommodate additional regularization steps within a DIP framework.

1623, **TITLE:** MRI Image Recovery using Damped Denoising Vector AMP

<https://doi.org/10.1109/ICASSP39728.2021.9415050>

**AUTHORS:** S. Sarkar, R. Ahmad and P. Schniter

**HIGHLIGHT:** Motivated by image recovery in magnetic resonance imaging (MRI), we propose a new approach to solving linear inverse problems based on iteratively calling a deep neural-network, sometimes referred to as plug-and-play recovery.

1624, **TITLE:** Overcoming Measurement Inconsistency In Deep Learning For Linear Inverse Problems: Applications In Medical Imaging

<https://doi.org/10.1109/ICASSP39728.2021.9414173>

**AUTHORS:** M. Vella and J. F. C. Mota

**HIGHLIGHT:** We show that such inconsistency, which can be critical in domains like medical imaging or defense, is intimately related to the generalization error. We then propose a framework that post-processes the output of DNNs with an optimization algorithm that enforces measurement consistency.

1625, **TITLE:** Scalable Reinforcement Learning For Routing In Ad-Hoc Networks Based On Physical-Layer Attributes

<https://doi.org/10.1109/ICASSP39728.2021.9413511>

**AUTHORS:** W. Cui and W. Yu

**HIGHLIGHT:** This work proposes a novel and scalable reinforcement learning approach for routing in ad-hoc wireless networks.

1626, **TITLE:** Blind Carbon Copy on Dirty Paper: Seamless Spectrum Underlay via Canonical Correlation Analysis

<https://doi.org/10.1109/ICASSP39728.2021.9414621>

**AUTHORS:** M. S. Ibrahim and N. D. Sidiropoulos

**HIGHLIGHT:** This work proposes a practical data-driven approach that allows a pair of secondary users to reliably communicate in underlay mode while keeping the interference at the primary receiver close to its noise floor.

1627, **TITLE:** An Actor-Critic Reinforcement Learning Approach to Minimum age of Information Scheduling in Energy Harvesting Networks

<https://doi.org/10.1109/ICASSP39728.2021.9415110>

**AUTHORS:** S. Leng and A. Yener

**HIGHLIGHT:** We consider the user scheduling problem over a communication session.

1628, **TITLE:** Moving Object Classification with a Sub-6 GHz Massive MIMO Array Using Real Data

<https://doi.org/10.1109/ICASSP39728.2021.9414952>

**AUTHORS:** B. R. Manoj, G. Tian, S. Gunnarsson, F. Tufvesson and E. G. Larsson

**HIGHLIGHT:** In this paper, we analyze classification of moving objects by employing machine learning on real data from a massive multi-input-multi-output (MIMO) system in an indoor environment.

- 1629, TITLE: Optimizing Coverage and Capacity in Cellular Networks using Machine Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9414155>  
AUTHORS: R. M. Dreifuerst et al.  
HIGHLIGHT: In this paper, we develop and compare two approaches for maximizing coverage and minimizing interference by jointly optimizing the transmit power and downtilt (elevation tilt) settings across sectors.
- 1630, TITLE: Unsupervised Learning for Asynchronous Resource Allocation In Ad-Hoc Wireless Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414181>  
AUTHORS: Z. Wang, M. Eisen and A. Ribeiro  
HIGHLIGHT: We consider optimal resource allocation problems under asynchronous wireless network setting.
- 1631, TITLE: Two-Stage Adaptive Pooling with RT-QPCR for Covid-19 Screening  
<https://doi.org/10.1109/ICASSP39728.2021.9413685>  
AUTHORS: A. Heidarzadeh and K. Narayanan  
HIGHLIGHT: We propose two-stage adaptive pooling schemes, 2-STAP and 2-STAMP, for detecting COVID-19 using real-time reverse transcription quantitative polymerase chain reaction (RT-qPCR) test kits.
- 1632, TITLE: Point of Care Image Analysis for COVID-19  
<https://doi.org/10.1109/ICASSP39728.2021.9413687>  
AUTHORS: D. Yaron et al.  
HIGHLIGHT: Here we train deep neural networks to significantly enhance the capability to detect, grade and monitor COVID-19 patients using CXRs and LUS.
- 1633, TITLE: An Improved Data Driven Dynamic SIRD Model for Predictive Monitoring of COVID-19  
<https://doi.org/10.1109/ICASSP39728.2021.9414762>  
AUTHORS: P. Singh, A. Singhal, B. Fatimah and A. Gupta  
HIGHLIGHT: In this work, we carry out the functional modeling of COVID-19 infection trends using two models: the Gaussian mixture model (GMM) and the composite logistic growth model (CLGM).
- 1634, TITLE: Leveraging A Multiple-Strain Model with Mutations in Analyzing the Spread of Covid-19  
<https://doi.org/10.1109/ICASSP39728.2021.9414595>  
AUTHORS: A. Sridhar, O. Yagan, R. Eletreby, S. A. Levin, J. B. Plotkin and H. V. Poor  
HIGHLIGHT: In this paper, we discuss how our recent work on a multiple-strain spreading model with mutations can help address some key questions concerning the spread of COVID-19.
- 1635, TITLE: Contact Tracing Enhances the Efficiency of Covid-19 Group Testing  
<https://doi.org/10.1109/ICASSP39728.2021.9414034>  
AUTHORS: R. Goenka, S. -J. Cao, C. -W. Wong, A. Rajwade and D. Baron  
HIGHLIGHT: In this paper, we use side information (SI) collected from contact tracing (CT) within nonadaptive/single-stage group testing algorithms.
- 1636, TITLE: Optimal Questionnaires for Screening of Strategic Agents  
<https://doi.org/10.1109/ICASSP39728.2021.9414345>  
AUTHORS: A. S. Vora and A. A. Kulkarni  
HIGHLIGHT: We investigate the problem of questioning travellers to classify them for further testing when the travellers are strategic or are unwilling to reveal their travel histories.
- 1637, TITLE: Exploring Visual-Audio Composition Alignment Network for Quality Fashion Retrieval in Video  
<https://doi.org/10.1109/ICASSP39728.2021.9413617>  
AUTHORS: Y. Zhang et al.  
HIGHLIGHT: In this paper, we present a novel Visual-Audio Composition Alignment Network (VACANet) to deal with quality fashion retrieval in video.
- 1638, TITLE: A Secure Searchable Image Retrieval Scheme with Correct Retrieval Identity  
<https://doi.org/10.1109/ICASSP39728.2021.9414024>  
AUTHORS: L. Wang and H. Yu  
HIGHLIGHT: Aiming at the issue of user privacy, we proposed a Secure searchable image retrieval scheme with correct retrieval identity.
- 1639, TITLE: Injecting Word Information with Multi-Level Word Adapter for Chinese Spoken Language Understanding

<https://doi.org/10.1109/ICASSP39728.2021.9413657>

AUTHORS: D. Teng, L. Qin, W. Che, S. Zhao and T. Liu

HIGHLIGHT: In this paper, we improve Chinese spoken language understanding (SLU) by injecting word information.

1640, TITLE: A Co-Interactive Transformer for Joint Slot Filling and Intent Detection

<https://doi.org/10.1109/ICASSP39728.2021.9414110>

AUTHORS: L. Qin, T. Liu, W. Che, B. Kang, S. Zhao and T. Liu

HIGHLIGHT: In this paper, we propose a Co-Interactive Transformer which considers the cross-impact between the two tasks.

1641, TITLE: Dual Metric Discriminator for Open Set Video Domain Adaptation

<https://doi.org/10.1109/ICASSP39728.2021.9413361>

AUTHORS: Y. Wang, X. Song, Y. Wang, P. Xu, R. Hu and H. Chai

HIGHLIGHT: In this paper, we propose a seminal framework, which involves spatial and temporal information to address OSVDA problem.

1642, TITLE: Cross-Domain Sentiment Classification with Contrastive Learning and Mutual Information Maximization

<https://doi.org/10.1109/ICASSP39728.2021.9414930>

AUTHORS: T. Li, X. Chen, S. Zhang, Z. Dong and K. Keutzer

HIGHLIGHT: In this work we propose a novel model for cross-domain sentiment classification - CLIM - Contrastive Learning with mutual Information Maximization, to explore the potential of contrastive learning for learning domain-invariant and task-discriminative features.

1643, TITLE: Low-Complexity Parameter Learning for OTFS Modulation Based Automotive Radar

<https://doi.org/10.1109/ICASSP39728.2021.9414107>

AUTHORS: C. Liu, S. Liu, Z. Mao, Y. Huang and H. Wang

HIGHLIGHT: In this work, we consider an OTFS modulation based automotive joint radar-communication system and focus on the design of low-complexity parameter estimation algorithm for radar targets.

1644, TITLE: Federated Dropout Learning for Hybrid Beamforming with Spatial Path Index Modulation in Multi-User Mmwave-Mimo Systems

<https://doi.org/10.1109/ICASSP39728.2021.9414120>

AUTHORS: A. M. Elbir, S. Coleri and K. V. Mishra

HIGHLIGHT: In this paper, we introduce model-based and model-free frameworks for beamformer design in multi-user SPIM-MIMO systems.

1645, TITLE: Information Decoding and SDR Implementation of DFRC Systems without Training Signals

<https://doi.org/10.1109/ICASSP39728.2021.9413379>

AUTHORS: D. M. Wong, B. K. Chalise, J. Metcalf and M. Amin

HIGHLIGHT: We propose a novel method for decoding information at the communication receiver without using training data, which is implemented using a software-defined radio (SDR).

1646, TITLE: A Low-Complexity MIMO Dual Function Radar Communication System via One-Bit Sampling

<https://doi.org/10.1109/ICASSP39728.2021.9414051>

AUTHORS: S. Zhu, F. Xi, S. Chen and A. Nehorai

HIGHLIGHT: In this paper, we propose to implement a low-complexity multiple input multiple output DFRC (MIMO-DFRC) system relying on the generalized spatial modulation (GSM) and the low-resolution sampling.

1647, TITLE: Learning to Select for Mimo Radar Based on Hybrid Analog-Digital Beamforming

<https://doi.org/10.1109/ICASSP39728.2021.9413904>

AUTHORS: Z. Xu, F. Liu, K. Diamantaras, C. Masouros and A. Petropulu

HIGHLIGHT: In this paper, we propose an energy-efficient radar beampattern design framework for Millimeter Wave (mmWave) massive multi-input multi-output (mMIMO) systems, equipped with a hybrid analog-digital (HAD) beamforming structure.

1648, TITLE: Word-Level ASL Recognition and Trigger Sign Detection with RF Sensors

<https://doi.org/10.1109/ICASSP39728.2021.9414063>

AUTHORS: M. M. Rahman et al.

HIGHLIGHT: Thus, this paper investigates RF sensing as an alternative sensing modality for ASL recognition to facilitate interactive devices and smart environments for the deaf and hard-of-hearing.

1649, TITLE: Hybrid Beamforming for Wideband OFDM Dual Function Radar Communications  
<https://doi.org/10.1109/ICASSP39728.2021.9413497>  
AUTHORS: Z. Cheng, J. He, S. Shi, Z. He and B. Liao  
HIGHLIGHT: This paper considers the hybrid beamforming design for wideband OFDM-DFRC system serving multiple users (MUs).

1650, TITLE: Bit Constrained Communication Receivers In Joint Radar Communications Systems  
<https://doi.org/10.1109/ICASSP39728.2021.9413979>  
AUTHORS: D. Ma, N. Shlezinger, T. Huang, Y. Liu and Y. C. Eldar  
HIGHLIGHT: In this paper, we design bit constrained communication receivers in dual-function systems, by considering hybrid analog/digital architectures and treating their operation as task-based quantization.

1651, TITLE: ICI-Aware Parameter Estimation for MIMO-Ofdm Radar via Apes Spatial Filtering  
<https://doi.org/10.1109/ICASSP39728.2021.9414537>  
AUTHORS: M. F. Keskin, H. Wymeersch and V. Koivunen  
HIGHLIGHT: We propose a novel three-stage delay-Doppler-angle estimation algorithm for a MIMO-OFDM radar in the presence of inter-carrier interference (ICI).

1652, TITLE: Joint Communications with FH-MIMO Radar Systems: An Extended Signaling Strategy  
<https://doi.org/10.1109/ICASSP39728.2021.9413462>  
AUTHORS: X. Wang, J. Xu, A. Hassanien and E. Aboutanios  
HIGHLIGHT: In this paper, we investigate the signaling strategy of communications embedding in frequency-hopping (FH) multiple input multiple output (MIMO) radar.

1653, TITLE: Full-Duplex Multifunction Transceiver with Joint Constant Envelope Transmission and Wideband Reception  
<https://doi.org/10.1109/ICASSP39728.2021.9413725>  
AUTHORS: J. Marin, M. Bernhardt and T. Riihonen  
HIGHLIGHT: This paper introduces and justifies a novel system concept that consists of full-duplex transceivers and uses a multifunction signal for simultaneous two-way communication, jamming and sensing tasks.

1654, TITLE: Waveform Design for the Joint MIMO Radar and Communications with Low Integrated Sidelobe Levels and Accurate Information Embedding  
<https://doi.org/10.1109/ICASSP39728.2021.9415119>  
AUTHORS: Y. Li, X. Wu and R. Tao  
HIGHLIGHT: In this paper, we focus on the multiple-waveform design for the joint multiple-input multiple-output radar and communications system, which aims to simultaneously attain low integrated sidelobe level (ISL) of waveforms and accurate fast-time modulation for information embedding (IE).

1655, TITLE: Ordered Reliability Bits Guessing Random Additive Noise Decoding  
<https://doi.org/10.1109/ICASSP39728.2021.9414615>  
AUTHORS: K. R. Duffy  
HIGHLIGHT: To that end, here we introduce a soft-detection variant of Guessing Random Additive Noise Decoding (GRAND) called Ordered Reliability Bits GRAND that can decode any moderate redundancy block-code.

1656, TITLE: Learned Decimation for Neural Belief Propagation Decoders : Invited Paper  
<https://doi.org/10.1109/ICASSP39728.2021.9414407>  
AUTHORS: A. Buchberger, C. H?ger, H. D. Pfister, L. Schmalen and A. G. i. Amat  
HIGHLIGHT: We introduce a two-stage decimation process to improve the performance of neural belief propagation (NBP), recently introduced by Nachmani et al., for short low-density parity-check (LDPC) codes.

1657, TITLE: ADMM-Based ML Decoding: from Theory to Practice  
<https://doi.org/10.1109/ICASSP39728.2021.9414728>  
AUTHORS: K. Kraft and N. Wehn  
HIGHLIGHT: In this paper, we investigate this approach with respect to its algorithmic and implementation-specific challenges.

1658, TITLE: Towards Practical Near-Maximum-Likelihood Decoding of Error-Correcting Codes: An Overview  
<https://doi.org/10.1109/ICASSP39728.2021.9414311>  
AUTHORS: T. Tonnellier, M. Hashemipour, N. Doan, W. J. Gross and A. Balatsoukas-Stimming  
HIGHLIGHT: This overview paper surveys recent progress in this emerging field by reviewing the GRAND algorithm, linear programming decoding, machine-learning aided decoding and the recursive projection-aggregation decoding algorithm.

- 1659, TITLE: High-Throughput VLSI Architecture for Soft-Decision Decoding with ORBGRAND  
<https://doi.org/10.1109/ICASSP39728.2021.9414908>  
AUTHORS: S. M. Abbas, T. Tonnellier, F. Ercan, M. Jaleldine and W. J. Gross  
HIGHLIGHT: This work reports the first-ever hardware architecture for ORBGRAND, which achieves an average throughput of up to 42.5 Gbps for a code length of 128 at an SNR of 10 dB.
- 1660, TITLE: Hardware Implementation of Iterative Projection-Aggregation Decoding of Reed-Muller Codes  
<https://doi.org/10.1109/ICASSP39728.2021.9414655>  
AUTHORS: M. Hashemipour-Nazari, K. Goossens and A. Balatsoukas-Stimming  
HIGHLIGHT: In this work, we present a simplification and a corresponding hardware architecture for hard-decision recursive projection-aggregation (RPA) decoding of Reed-Muller (RM) codes.
- 1661, TITLE: m-Activity: Accurate and Real-Time Human Activity Recognition Via Millimeter Wave Radar  
<https://doi.org/10.1109/ICASSP39728.2021.9414686>  
AUTHORS: Y. Wang, H. Liu, K. Cui, A. Zhou, W. Li and H. Ma  
HIGHLIGHT: In this paper, we propose m-Activity, which can realize HAR while reducing noise caused by environmental multi-path effects, and operate fluently at runtime.
- 1662, TITLE: Pushing the Limit of Phase Offset for Contactless Sensing Using Commodity Wifi  
<https://doi.org/10.1109/ICASSP39728.2021.9414926>  
AUTHORS: D. Zhang, X. Li and Y. Chen  
HIGHLIGHT: In this paper, we propose BreathTrack2.0, a contactless breath tracking system based on commodity WiFi devices.
- 1663, TITLE: Noncontact Heartbeat Detection by Viterbi Algorithm with Fusion of Beat-Beat Interval and Deep Learning-Driven Branch Metrics  
<https://doi.org/10.1109/ICASSP39728.2021.9413401>  
AUTHORS: K. Yamamoto and T. Ohtsuki  
HIGHLIGHT: In this paper, to detect heartbeat with a high accuracy, we propose a Doppler radar-based heartbeat detection method by the Viterbi algorithm with a fusion of Beat-Beat Interval (BBI) and deep learning-driven Branch Metrics (BM).
- 1664, TITLE: Typingwristband: A Human Slight Motion Sensing System Based on Vibration Detection  
<https://doi.org/10.1109/ICASSP39728.2021.9414751>  
AUTHORS: S. Cheng, J. Yan, J. Li and J. Liu  
HIGHLIGHT: In this paper, we focus on the problem of how to detect the human's typing motion, and designed a new system, named as Typing Wristband, to obtain the vibration of wrist using piezoelectric transducer (PZT).
- 1665, TITLE: Movement Detection Using A Reciprocal Received Signal Strength Model  
<https://doi.org/10.1109/ICASSP39728.2021.9414717>  
AUTHORS: O. Kaltiokallio and H. Yigitler  
HIGHLIGHT: A novel reciprocal signal strength model is presented, and an energy detector is developed.
- 1666, TITLE: Deep Convolutional Gaussian Processes for Mmwave Outdoor Localization  
<https://doi.org/10.1109/ICASSP39728.2021.9414388>  
AUTHORS: X. Wang, M. Patil, C. Yang, S. Mao and P. A. Patel  
HIGHLIGHT: In this paper, we propose a deep convolutional Gaussian process (DCGP) based regression approach to achieve high robustness for fingerprinting-based mmWave outdoor localization, which exploits the convolutional structure for deep Gaussian process to allow uncertainty estimation on location predictions.
- 1667, TITLE: Exploring Automatic COVID-19 Diagnosis via Voice and Symptoms from Crowdsourced Data  
<https://doi.org/10.1109/ICASSP39728.2021.9414576>  
AUTHORS: J. Han et al.  
HIGHLIGHT: In this paper, we propose a voice-based framework to automatically detect individuals who have tested positive for COVID-19.
- 1668, TITLE: Coughwatch: Real-World Cough Detection using Smartwatches  
<https://doi.org/10.1109/ICASSP39728.2021.9414881>  
AUTHORS: D. Liaqat et al.



**HIGHLIGHT:** In this work we propose CoughWatch, a lightweight cough detector for smartwatches that uses audio and movement data for in-the-wild cough detection.

1669, **TITLE:** Acoustic and Linguistic Analyses to Assess Early-Onset and Genetic Alzheimer's Disease  
<https://doi.org/10.1109/ICASSP39728.2021.9414009>

**AUTHORS:** P. A. Pérez-Toro et al.

**HIGHLIGHT:** This study proposes the use of acoustic and linguistic methods to extract features from speech recordings and their transcriptions to discriminate people with conditions related to the Paisa mutation.

1670, **TITLE:** A Noise-Robust Signal Processing Strategy for Cochlear Implants Using Neural Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9413452>

**AUTHORS:** N. Zheng, Y. Shi, Y. Kang and Q. Meng

**HIGHLIGHT:** This paper presents a noise-robust signal processing strategy to deal with this problem.

1671, **TITLE:** Context-Aware Speech Stress Detection in Hospital Workers Using Bi-LSTM Classifiers  
<https://doi.org/10.1109/ICASSP39728.2021.9414666>

**AUTHORS:** A. Gaballah, A. Tiwari, S. Narayanan and T. H. Falk

**HIGHLIGHT:** In this paper, we present a context-aware speech-based system for stress detection.

1672, **TITLE:** Unsupervised Heart Abnormality Detection Based on Phonocardiogram Analysis with Beta Variational Auto-Encoders  
<https://doi.org/10.1109/ICASSP39728.2021.9414165>

**AUTHORS:** S. Li, K. Tian and R. Wang

**HIGHLIGHT:** This paper proposes a method of unsupervised PCG analysis that uses beta variational auto-encoder (VAE) to model the normal PCG signals.

1673, **TITLE:** Compressing Deep Neural Networks for Efficient Speech Enhancement  
<https://doi.org/10.1109/ICASSP39728.2021.9413536>

**AUTHORS:** K. Tan and D. Wang

**HIGHLIGHT:** In order to address this problem, we propose a model compression pipeline to reduce DNN size for speech enhancement, which is based on three kinds of techniques: sparse regularization, iterative pruning and clustering-based quantization.

1674, **TITLE:** Improved Mask-CTC for Non-Autoregressive End-to-End ASR  
<https://doi.org/10.1109/ICASSP39728.2021.9414198>

**AUTHORS:** Y. Higuchi, H. Inaguma, S. Watanabe, T. Ogawa and T. Kobayashi

**HIGHLIGHT:** To boost the performance of Mask-CTC, we first propose to enhance the encoder network architecture by employing a recently proposed architecture called Conformer. Next, we propose new training and decoding methods by introducing auxiliary objective to predict the length of a partial target sequence, which allows the model to delete or insert tokens during inference.

1675, **TITLE:** Memory-Efficient Speech Recognition on Smart Devices  
<https://doi.org/10.1109/ICASSP39728.2021.9414502>

**AUTHORS:** G. Venkatesh et al.

**HIGHLIGHT:** We address transducer model's memory access concerns by optimizing their model architecture and designing novel recurrent cell designs.

1676, **TITLE:** Expediting discovery in Neural Architecture Search by Combining Learning with Planning  
<https://doi.org/10.1109/ICASSP39728.2021.9413547>

**AUTHORS:** F. S. Fard and V. Singh Tomar

**HIGHLIGHT:** Time to discover optimal structures is a key concern in many AML solutions including NASIL. Here, we proposed an extended version called "GNASIL" to speed up the process.

1677, **TITLE:** Specialized Embedding Approximation for Edge Intelligence: A Case Study in Urban Sound Classification  
<https://doi.org/10.1109/ICASSP39728.2021.9414287>

**AUTHORS:** S. Srivastava, D. Roy, M. Cartwright, J. P. Bello and A. Arora

**HIGHLIGHT:** Hence, we introduce Specialized Embedding Approximation (SEA) to train a student featurizer to approximate the teacher's embedding manifold for a given target domain.

1678, **TITLE:** Light-TTS: Lightweight Multi-Speaker Multi-Lingual Text-to-Speech  
<https://doi.org/10.1109/ICASSP39728.2021.9414400>

**AUTHORS:** S. Li, B. Ouyang, L. Li and Q. Hong

**HIGHLIGHT:** In this paper, we propose a new lightweight multi-speaker multi-lingual speech synthesis system, named LightTTS, which can quickly synthesize the Chinese, English or code-switch speech of multiple speakers in a non-autoregressive generation manner using only one model.

1679, **TITLE:** Efficient Long Periodic Binary Sequence Designs for Automotive Radar

<https://doi.org/10.1109/ICASSP39728.2021.9413747>

**AUTHORS:** Y. Chen, R. Lin and J. Li

**HIGHLIGHT:** We consider an FFT-based algorithm for the efficient designs of long periodic binary sequences with arbitrary period lengths and ample diversity.

1680, **TITLE:** Joint Localization and Predictive Beamforming in Vehicular Networks: Power Allocation Beyond Water-Filling

<https://doi.org/10.1109/ICASSP39728.2021.9414871>

**AUTHORS:** F. Liu and C. Masouros

**HIGHLIGHT:** This paper explores tailored power allocation (PA) for dual functional radar-communication (DFRC) in the vehicle-to-infrastructure (V2I) network, where a road side unit (RSU) provides both localization and communication services to multiple vehicles.

1681, **TITLE:** A New Automotive Radar 4D Point Clouds Detector by Using Deep Learning

<https://doi.org/10.1109/ICASSP39728.2021.9413682>

**AUTHORS:** Y. Cheng, J. Su, H. Chen and Y. Liu

**HIGHLIGHT:** Therefore, in this paper, we propose a new automotive radar detector based on deep learning using the spatial distribution feature of the real targets, in order to improve the performance of automotive radar detector in the real-world driving scene.

1682, **TITLE:** Enhanced Automotive Target Detection through Radar and Communications Sensor Fusion

<https://doi.org/10.1109/ICASSP39728.2021.9415051>

**AUTHORS:** S. H. Dokhanchi, B. Shankar Mysore, K. V. Mishra and B. Ottersten

**HIGHLIGHT:** A sensor fusion algorithm is proposed to benefit from the information from radar and communication to improve the final range estimates.

1683, **TITLE:** Extended Object Tracking With Automotive Radar Using B-Spline Chained Ellipses Model

<https://doi.org/10.1109/ICASSP39728.2021.9415080>

**AUTHORS:** G. Yao, P. Wang, K. Berntorp, H. Mansour, P. Boufounos and P. V. Orlik

**HIGHLIGHT:** This paper introduces a B-spline chained ellipses model representation for extended object tracking (EOT) using high-resolution automotive radar measurements.

1684, **TITLE:** Four-Dimensional High-Resolution Automotive Radar Imaging Exploiting Joint Sparse-Frequency and Sparse-Array Design

<https://doi.org/10.1109/ICASSP39728.2021.9413875>

**AUTHORS:** S. Sun and Y. D. Zhang

**HIGHLIGHT:** We propose a novel automotive radar imaging technique to provide high-resolution information in four dimensions, i.e., range, Doppler, azimuth, and elevation, by exploiting a joint sparsity design in frequency spectrum and array configurations.

1685, **TITLE:** An Empirical Study of Visual Features for DNN Based Audio-Visual Speech Enhancement in Multi-Talker Environments

<https://doi.org/10.1109/ICASSP39728.2021.9414000>

**AUTHORS:** S. S. Shetu, S. Chakrabarty and E. A. P. Habets

**HIGHLIGHT:** In this work, we perform an empirical study of the most commonly used visual features for DNN based AVSE, the pre-processing requirements for each of these features, and investigate their influence on the performance.

1686, **TITLE:** On The Role of Visual Cues in Audiovisual Speech Enhancement

<https://doi.org/10.1109/ICASSP39728.2021.9414263>

**AUTHORS:** Z. Aldeneh et al.

**HIGHLIGHT:** We present an introspection of an audiovisual speech enhancement model.

1687, **TITLE:** Convolutional Transfer Function Invariant SDR Training Criteria for Multi-Channel Reverberant Speech Separation

<https://doi.org/10.1109/ICASSP39728.2021.9414661>

**AUTHORS:** C. Boeddeker et al.

**HIGHLIGHT:** For the objective we propose to use a convolutive transfer function invariant Signal-to-Distortion Ratio (CI-SDR) based loss.

1688, **TITLE:** Directional ASR: A New Paradigm for E2E Multi-Speaker Speech Recognition with Source Localization  
<https://doi.org/10.1109/ICASSP39728.2021.9414243>

**AUTHORS:** A. S. Subramanian et al.

**HIGHLIGHT:** This paper proposes a new paradigm for handling far-field multi-speaker data in an end-to-end (E2E) neural network manner, called directional automatic speech recognition (D-ASR), which explicitly models source speaker locations.

1689, **TITLE:** Communication-Cost Aware Microphone Selection for Neural Speech Enhancement with Ad-Hoc Microphone Arrays  
<https://doi.org/10.1109/ICASSP39728.2021.9414775>

**AUTHORS:** J. Casebeer, J. Kaikau and P. Smaragdis

**HIGHLIGHT:** In this paper, we present a method for jointly-learning a microphone selection mechanism and a speech enhancement network for multi-channel speech enhancement with an ad-hoc microphone array.

1690, **TITLE:** Deep Multi-Frame MVDR Filtering for Single-Microphone Speech Enhancement  
<https://doi.org/10.1109/ICASSP39728.2021.9413775>

**AUTHORS:** M. Tammen and S. Doclo

**HIGHLIGHT:** Aiming at merging the speech enhancement potential of the MFMVDR filter and the estimation capability of temporal convolutional networks (TCNs), in this paper we propose to embed the MFMVDR filter within a deep learning framework.

1691, **TITLE:** Compressive Wideband Spectrum Sensing and Carrier Frequency Estimation with Unknown MIMO Channels  
<https://doi.org/10.1109/ICASSP39728.2021.9413839>

**AUTHORS:** H. Wang, J. Wang, J. Fang and H. Li

**HIGHLIGHT:** We consider the problem of joint wideband spectrum sensing and carrier frequency estimation in a sub-Nyquist sampling framework.

1692, **TITLE:** Joint Optimization of Spectrally Co-Existing Multi-Carrier Radar and Communication Systems in Cluttered Environments  
<https://doi.org/10.1109/ICASSP39728.2021.9413995>

**AUTHORS:** F. Wang, H. Li and B. Himed

**HIGHLIGHT:** We propose a non-alternating approach to jointly optimize the radar and communication transmission power allocated to each sub-carrier.

1693, **TITLE:** Target Detection in Frequency Hopping MIMO Dual-Function Radar-Communication Systems  
<https://doi.org/10.1109/ICASSP39728.2021.9413852>

**AUTHORS:** I. P. Eedara, M. G. Amin and G. A. Fabrizio

**HIGHLIGHT:** We consider a multiple-input multiple-output (MIMO) dual function radar communication (DFRC) system employing frequency hopping (FH) radar waveforms.

1694, **TITLE:** Asymptotic Distribution of Generalized Likelihood Ratio Test Under Model Misspecification With Application to Cooperative Radar-Communications  
<https://doi.org/10.1109/ICASSP39728.2021.9414036>

**AUTHORS:** A. S. Bondre and C. D. Richmond

**HIGHLIGHT:** The goal of this paper is to develop an expression for the asymptotic distribution of the generalized likelihood ratio test (GLRT) statistic under model misspecification, that is when the assumed data model is different from the true model.

1695, **TITLE:** Online Antenna Selection for Enhanced DOA Estimation  
<https://doi.org/10.1109/ICASSP39728.2021.9414591>

**AUTHORS:** E. Aboutanios, H. Nosrati and X. Wang

**HIGHLIGHT:** We formulate the changing DOA estimation problem as a game in the context of online convex optimization, and employ a gradient-based technique that makes a move at each step in order to minimize the total loss after T steps.

1696, **TITLE:** Designing Random FM Radar Waveforms with Compact Spectrum  
<https://doi.org/10.1109/ICASSP39728.2021.9414020>

**AUTHORS:** C. A. Mohr and S. D. Blunt

**HIGHLIGHT:** Here we examine the impact of using the family of super-Gaussian spectra to serve as alternative design templates.

- 1697, TITLE: Collaborative Inference via Ensembles on the Edge  
<https://doi.org/10.1109/ICASSP39728.2021.9414740>  
AUTHORS: N. Shlezinger, E. Farhan, H. Morgenstern and Y. C. Eldar  
HIGHLIGHT: Here we propose a framework for facilitating the application of DNNs on the edge in a manner which allows multiple users to collaborate during inference in order to improve their prediction accuracy.
- 1698, TITLE: Allocating DNN Layers Computation Between Front-End Devices and The Cloud Server for Video Big Data Processing  
<https://doi.org/10.1109/ICASSP39728.2021.9413867>  
AUTHORS: P. Xing, X. Liu, P. Peng, T. Huang and Y. Tian  
HIGHLIGHT: This paper proposes a computation allocation algorithm of DNN between the front-end devices and the cloud server.
- 1699, TITLE: Branchy-GNN: A Device-Edge Co-Inference Framework for Efficient Point Cloud Processing  
<https://doi.org/10.1109/ICASSP39728.2021.9414831>  
AUTHORS: J. Shao, H. Zhang, Y. Mao and J. Zhang  
HIGHLIGHT: Built upon the emerging paradigm of device-edge co-inference, where an edge device extracts and transmits the intermediate feature to an edge server for further processing, we propose Branchy-GNN for efficient graph neural network (GNN) based point cloud processing by leveraging edge computing platforms.
- 1700, TITLE: Collaborative Intelligence: Challenges and Opportunities  
<https://doi.org/10.1109/ICASSP39728.2021.9413943>  
AUTHORS: I. V. Bajic, W. Lin and Y. Tian  
HIGHLIGHT: This paper presents an overview of the emerging area of collaborative intelligence (CI).
- 1701, TITLE: Latent Space Motion Analysis for Collaborative Intelligence  
<https://doi.org/10.1109/ICASSP39728.2021.9413603>  
AUTHORS: M. Ulhaq and I. V. Bajic  
HIGHLIGHT: By analyzing the effect of common DNN operations on optical flow, we show that the motion present in each channel of a feature tensor is approximately equal to the scaled version of the input motion.
- 1702, TITLE: Teacher-Student Learning With Multi-Granularity Constraint Towards Compact Facial Feature Representation  
<https://doi.org/10.1109/ICASSP39728.2021.9413506>  
AUTHORS: S. Wang, S. Wang, W. Yang, X. Zhang, S. Wang and S. Ma  
HIGHLIGHT: In this paper, we propose a novel end-to-end feature compression scheme by leveraging the representation and learning capability of deep neural networks, towards intelligent front-end equipped analysis with promising accuracy and efficiency.
- 1703, TITLE: Discriminability of Single-Layer Graph Neural Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414583>  
AUTHORS: S. Pfrommer, A. Ribeiro and F. Gama  
HIGHLIGHT: We focus on the property of discriminability and establish conditions under which the inclusion of pointwise nonlinearities to a stable graph filter bank leads to an increased discriminative capacity for high-eigenvalue content.
- 1704, TITLE: On The Stability of Graph Convolutional Neural Networks Under Edge Rewiring  
<https://doi.org/10.1109/ICASSP39728.2021.9413474>  
AUTHORS: H. Kenlay, D. Thanos and X. Dong  
HIGHLIGHT: In this work, we develop an interpretable upper bound elucidating that graph neural networks are stable to rewiring between high degree nodes.
- 1705, TITLE: Geometric Scattering Attention Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414557>  
AUTHORS: Y. Min, F. Wenkel and G. Wolf  
HIGHLIGHT: Here, we introduce a new attention-based architecture to produce adaptive task-driven node representations by implicitly learning node-wise weights for combining multiple scattering and GCN channels in the network.
- 1706, TITLE: Ego-GNNs: Exploiting Ego Structures in Graph Neural Networks  
<https://doi.org/10.1109/ICASSP39728.2021.9414015>  
AUTHORS: D. Sandfelder, P. Vijayan and W. L. Hamilton  
HIGHLIGHT: In this work, we propose to augment the GNN message-passing operations with information de-fined on ego graphs (i.e., the induced subgraph surrounding each node).

- 1707, TITLE: Learning the Relevant Substructures for Tasks on Graph Data  
<https://doi.org/10.1109/ICASSP39728.2021.9414377>  
AUTHORS: L. Chen, Z. Chen and J. Bruna  
HIGHLIGHT: In this work, we demonstrate that LRP models can be used on challenging graph classification tasks to provide both state-of-the-art performance and interpretability, through the detection of the relevant substructures used by the network to make its decisions.
- 1708, TITLE: A Short Tutorial on The Weisfeiler-Lehman Test And Its Variants  
<https://doi.org/10.1109/ICASSP39728.2021.9413523>  
AUTHORS: N. T. Huang and S. Villar  
HIGHLIGHT: The goal of this short note is pedagogical and practical: We explain the differences between the WL and folklore-WL formulations, with pointers to existing discussions in the literature.
- 1709, TITLE: Hybrid Model for Network Anomaly Detection with Gradient Boosting Decision Trees and Tabtransformer  
<https://doi.org/10.1109/ICASSP39728.2021.9414766>  
AUTHORS: X. Xu and X. Zheng  
HIGHLIGHT: In this pa-per, we present our solution for the ICASSP 2021 Network Anomaly Detection Challenge (NAD) challenge.
- 1710, TITLE: Voting-Based Ensemble Model for Network Anomaly Detection  
<https://doi.org/10.1109/ICASSP39728.2021.9414532>  
AUTHORS: T. -H. Yang, Y. -T. Lin, C. -L. Wu and C. -Y. Wang  
HIGHLIGHT: In this work, we propose a machine learning framework based on XGBoost and deep neural networks to classify normal traffic and anomalous traffic.
- 1711, TITLE: An Accuracy Network Anomaly Detection Method Based on Ensemble Model  
<https://doi.org/10.1109/ICASSP39728.2021.9414675>  
AUTHORS: F. Liu, X. Li, W. Xiong, H. Jiang and G. Xie  
HIGHLIGHT: This paper introduces an ensemble model, which is a powerful technique to increase accuracy on network anomaly detection.
- 1712, TITLE: Fden: Mining Effective Information of Features in Detecting Network Anomalies  
<https://doi.org/10.1109/ICASSP39728.2021.9415099>  
AUTHORS: B. Li et al.  
HIGHLIGHT: In this paper, we propose a novel method to effectively leverage the features in detecting network anomalies, named FDEn, consisting of flow-based Feature Derivation (FD) and prior knowledge incorporated Ensemble models (Enpk).
- 1713, TITLE: Multi-Scale Residual Network for Covid-19 Diagnosis Using Ct-Scans  
<https://doi.org/10.1109/ICASSP39728.2021.9414426>  
AUTHORS: P. Garg, R. Ranjan, K. Upadhyay, M. Agrawal and D. Deepak  
HIGHLIGHT: This paper proposes a three-level approach to separate the cases of COVID-19, pneumonia from normal patients using chest CT scans.
- 1714, TITLE: Diagnosing Covid-19 from CT Images Based on an Ensemble Learning Framework  
<https://doi.org/10.1109/ICASSP39728.2021.9413707>  
AUTHORS: B. Li, Q. Zhang, Y. Song, Z. Zhao, Z. Meng and F. Su  
HIGHLIGHT: In this paper, we propose a novel ensemble learning framework to solve this problem.
- 1715, TITLE: CNR-IEMN: A Deep Learning Based Approach to Recognise Covid-19 from CT-Scan  
<https://doi.org/10.1109/ICASSP39728.2021.9414185>  
AUTHORS: F. Bougourzi, R. Contino, C. Distanto and A. Taleb-Ahmed  
HIGHLIGHT: In this paper, we proposed deep learning based approach to recognize the Covid-19 infection from the CT-scans.
- 1716, TITLE: Covid-19 Diagnostic Using 3d Deep Transfer Learning for Classification of Volumetric Computerised Tomography Chest Scans  
<https://doi.org/10.1109/ICASSP39728.2021.9414947>  
AUTHORS: S. Xue and C. Abhayaratne  
HIGHLIGHT: This paper proposes COVID-19 diagnosis based on analysis of Computerised tomography (CT) chest scans.

1717, TITLE: A Multi-Stage Progressive Learning Strategy for Covid-19 Diagnosis Using Chest Computed Tomography with Imbalanced Data  
<https://doi.org/10.1109/ICASSP39728.2021.9414745>

AUTHORS: Z. Yang, Y. Hou, Z. Chen, L. Zhang and J. Chen

HIGHLIGHT: In this paper, a multi-stage progressive learning strategy is investigated to train classifiers for COVID-19 Diagnosis using imbalanced Chest Computed Tomography Data acquired from patients infected with COVID-19 Pneumonia, Community Acquired Pneumonia (CAP) and from normal healthy subjects.

1718, TITLE: Detecting Covid-19 and Community Acquired Pneumonia Using Chest CT Scan Images With Deep Learning  
<https://doi.org/10.1109/ICASSP39728.2021.9414007>

AUTHORS: S. Chaudhary, S. Sadbhawna, V. Jakhetiya, B. N. Subudhi, U. Baid and S. C. Guntuku

HIGHLIGHT: We propose a two-stage Convolutional Neural Network (CNN) based classification framework for detecting COVID-19 and Community Acquired Pneumonia (CAP) using the chest Computed Tomography (CT) scan images.

1719, TITLE: Investigating on Incorporating Pretrained and Learnable Speaker Representations for Multi-Speaker Multi-Style Text-to-Speech  
<https://doi.org/10.1109/ICASSP39728.2021.9413880>

AUTHORS: C. -M. Chien, J. -H. Lin, C. -y. Huang, P. -c. Hsu and H. -y. Lee

HIGHLIGHT: In this work, we investigate different speaker representations and proposed to integrate pretrained and learnable speaker representations.

1720, TITLE: The Thinkit System for Icassp2021 M2voc Challenge  
<https://doi.org/10.1109/ICASSP39728.2021.9413669>

AUTHORS: Z. Shang, H. Zhang, Z. Chen, B. Zhou and P. Zhang

HIGHLIGHT: In this paper, we introduce the low resource text-to-speech system from the ThinkIT team submitted to Multi-Speaker Multi-Style Voice Cloning Challenge (M2VoC).

1721, TITLE: Dian: Duration Informed Auto-Regressive Network for Voice Cloning  
<https://doi.org/10.1109/ICASSP39728.2021.9414727>

AUTHORS: W. Song et al.

HIGHLIGHT: In this paper, we propose a novel end-to-end speech synthesis approach, Duration Informed Auto-regressive Network (DIAN), which consists of an acoustic model and a separate duration model.

1722, TITLE: Prosody and Voice Factorization for Few-Shot Speaker Adaptation in the Challenge M2voc 2021  
<https://doi.org/10.1109/ICASSP39728.2021.9414427>

AUTHORS: T. Wang et al.

HIGHLIGHT: To prevent the model from overfitting, this paper proposes a novel speaker adaptation framework that decomposes the prosody and voice characteristics in the end-to-end model.

1723, TITLE: The Huya Multi-Speaker and Multi-Style Speech Synthesis System for M2voc Challenge 2020  
<https://doi.org/10.1109/ICASSP39728.2021.9414943>

AUTHORS: J. Wang et al.

HIGHLIGHT: In this paper, we propose the Huya multi-speaker and multi-style speech synthesis system which is based on DurlAN and HiFi-GAN to generate high-fidelity speech even under low-resource condition.

1724, TITLE: The Multi-Speaker Multi-Style Voice Cloning Challenge 2021  
<https://doi.org/10.1109/ICASSP39728.2021.9414001>

AUTHORS: Q. Xie et al.

HIGHLIGHT: In this paper, we present a detailed explanation on the tasks and data used in the challenge, followed by a summary of submitted systems and evaluation results.