ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2025)

Hyderabad, India
6-11 April 2025

Pages 1-485

IEEE Catalog Number: CFP25ICA-POD
ISBN: 979-8-3503-6875-8

Copyright © 2025 by the Institute of Electrical and Electronics Engineers, Inc.
All Rights Reserved

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

****** This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP25ICA-POD
ISBN (Print-On-Demand):	979-8-3503-6875-8
ISBN (Online):	979-8-3503-6874-1
ISSN:	1520-6149

Additional Copies of This Publication Are Available From:

Curran Associates, Inc
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: (845) 758-0400
Fax: (845) 758-2633
E-mail: curran@proceedings.com
Web: www.proceedings.com

TABLE OF CONTENTS

HypCAD: Geometry-Enhanced Hyperbolic Contrastive Learning for CAD Model Retrieval 1
Adam Misik, Driton Salihu, Xin Su, Heike Brock, Eckehard Steinbach

Attention Augmented Structure-Centric Bias Mitigation with Feature Disentanglement 6
Xuege Hou, Yali Li, Shengjin Wang

Context-Guided Active Domain Adaptation for Blended Target Domain 11
Yuwu Lu, Yihan Yang

Enhancing Graph-Based Fraud Detection by Adversarial Confidence Reweighting 16
Jianqi Gao, Jian Cao, Shiyou Qian, Wei Guan

Attention Disentanglement for Semantic Diffusion Modeling in Text-To-Image Generation..... 21
Hsiang-Chun Yu, Jen-Tsung Chien

Weighted Density for the Win: Accurate Subspace Density Clustering 26
Maixuan Peng, Yuyang Wu, Yang Lu, Mengke Li, Yiqun Zhang, Yiu-Ming Cheung

FedTLU: Federated Learning with Targeted Layer Updates 31
Jong-Ik Park, Carlee Joe-Wong

Multiple Consistency-Guided Test-Time Adaptation for Contrastive Audio-Language Models with Unlabeled Audio..... 36
Gongyu Chen, Haomin Zhang, Chaofan Ding, Zihao Chen, Xinhan Di

Animation Anycolor: Enhancing Line Drawing Colorization with Keypoint Matching..... 41
Liyao Wang, Zuzeng Lin, Danni Wu, Zihao Yu, Suzhe Zhang, Zixian Wu, Feng Wang

In-Context Multitask Learning for Few-Shot Fine-Tuning of Large Language Models in Traditional Chinese Medicine Tongue Diagnosis 46
Changzeng Fu, Zelin Fu, Shaojun Yan, Xiaoyong Lyv, Yuliang Zhao

Identification and Correction of Permutation Errors in Compressed Sensing-Based Group Testing 51
Shuvayan Banerjee, Sudhansh Peddabomma, Radhendushka Srivastava, James Saunderson, Ajit Rajwade

RestorMamba: An Enhanced Synergistic State Space Model for Image Restoration..... 56
Zeyu Wang, Chen Li, Huiying Xu, Xinzhong Zhu, Xiao Huang, Hongbo Li

Improving Knowledge Base Question Answering Via Retrieval Enhancement and Stepwise Reasoning 61
Dian Huang, Jianqi Gao, Xiangfeng Luo, Hao Wu

MambaInst: Lightweight State Space Model for Real-Time Instance Segmentation 66
Zeyu Wang, Chen Li, Huiying Xu, Xinzhong Zhu, Xiao Huang, Hongbo Li

Robust Qualitative Data Clustering Via Learnable Multi-Metric Space Fusion..... 71
Sen Feng, Mingjie Zhao, Zhanpei Huang, Yuzhu Ji, Yiqun Zhang, Yiu-Ming Cheung

Improving Dialect Identification in Indian Languages Using Multimodal Features from Dialect Informed ASR..... 76
Amartyaveer, Saurabh Kumar, Sumit Sharma, Sathvik Udupa, Sandhya Badiger, Abhayjeet Singh, Deekshitha G, Jesuraja Bandekar, Savitha Murthy, Prasanta Kumar Ghosh

Gram: A Large-Scale General EEG Model for Raw Data Classification and Restoration Tasks 81
Ziyi Li, Wei-Long Zheng, Bao-Liang Lu

Content-Aware Dynamic Superpixel Segmentation 86
Tingyu Zhao, Bo Peng, Zhenguang Zhang, Daipeng Yang, Xi Wu

Online Optimization of Offloading Video Analytics Tasks to Multiple Edges for Accuracy Maximization..... 91
Yu Liang, Sheng Zhang, Jie Wu

A Risk Prediction Model for Real Estate Corporations Using High-Target Semantic BERT and Improved GRU 96
Xiaoyan Ma, Peng Zhu, Qinyuan Liu, Zidong Wang

Multi-Domain Fusion Network for Underwater Image Enhancement 101
Junbin Zhuang, Jiajia Zhou, Yan Zheng, Yasheng Chang, Suleman Mazhar

KARST: Multi-Kernel Kronecker Adaptation with Re-Scaling Transmission for Visual Classification 106
Yue Zhu, Haiwen Diao, Shang Gao, Long Chen, Huchuan Lu

Meta-MMD Fusion: Enhancing Cross-Subject Motor Imagery Classification 111
Minghui Chen, Chao Qu, Jiahui Pan

A MoE Multimodal Graph Attention Network Framework for Multimodal Emotion Recognition 116
Chengwen Zhang, Yaohui Liu, Bo Cheng

Efficient and Expandable Token-Level Approach for Multi-Domain Sensitive Information Classification 121
Hongyi Li, Jiawei Ye, Jie Wu, Lijun Zu

Follow-Your-MultiPose: Tuning-Free Multi-Character Text-To-Video Generation Via Pose Guidance..... 126
Beiyuan Zhang, Yue Ma, Chunlei Fu, Xinyang Song, Zhenan Sun, Ziqiang Li

Role-Specific Reward Design with Large Language Model for StarCraft II 131
Sijia Li, Haonan Lou, Xu Zhang, Xin Zeng, Zhixuan Shen, Tianrui Li

WebSurfer: Enhancing LLM Agents with Web-Wise Feedback for Web Navigation 136
Die Hu, Jingguo Ge, Weitao Tang, Guoyi Li, Liangxiong Li, Bingzhen Wu

Exploiting Foundation Models for Label-Efficient Few-Shot Learning Via Feature Coupling: A Case Study of Cardiac CT Segmentation 141
Wei Chen, Chen Li, Wenjuan Zhou, Yuhang Li, Tianhang Guo, Yuhua Tang

Filtering Resistant Large Language Model Watermarking Via Style Injection 146
Zhaojun Guo, Guobiao Li, Junqiang Huang, Xinpeng Zhang, Zhenxing Qian, Sheng Li

CSS: Overcoming Pose and Scene Challenges in Crowd-Sourced 3D Gaussian Splatting..... 151
Runze Chen, Mingyu Xiao, Haiyong Luo, Fang Zhao, Fan Wu, Hao Xiong, Qi Liu, Meng Song

PDCE: Patch-Wise Dynamic Curve Estimation for Low-Light Image Enhancement..... 156
Ruiyuan Chen, Zhixin Li, Han Zeng, Yifan Liu, Tao He, Tiecheng Song

AnimateSketches: Animate Sketches with Instance-Aware Mask..... 161
Haoge Deng, Xin Dai, Jijin Hu, Yonggang Qi

Detecting OOD Samples Via Optimal Transport Scoring Function 166
Heng Gao, Zhuolin He, Jian Pu

PGD-Imp: Rethinking and Unleashing Potential of Classic PGD with Dual Strategies for Imperceptible Adversarial Attacks..... 171
Jin Li, Zitong Yu, Ziqiang He, Z. Jane Wang, Xiangui Kang

Easy, Interpretable, Effective: openSMILE for Voice Deepfake Detection..... 176
Octavian Pascu, Dan Oneaţă, Horia Cucu, Nicolas Müller

A Divide-And-Conquer Approach for Sparse Recovery in High Dimensions 181
Aron Bevelander, Kim Batselier, Nitin Jonathan Myers

Imitating Human Selective Attention Using Dual Policy Network for Scanpath Prediction..... 186
Kepei Zhang, Ge Tong, Xuetao Zhang

GPPT: Gaussian Process-Infused Prompt Tuning for Vision-Language Models..... 191
Shijing Si, Haixia Sun, Jiawen Gu

AS-Net: Adaptive Style-Aware Network for Handwritten Text Generation..... 196
Yiming Wang, Hongxi Wei, Heng Wang

SmartExp: An Adaptive Data Expansion Strategy for Improving Handwritten Text Recognition..... 201
Yiming Wang, Hongxi Wei, Shiwen Sun

Promoting PLM Fine-Tuning Through Consistency Adversarial Training..... 206
Jianqi Gao, Jian Cao, Jinghua Tang

Algorithm Design for Continual Learning in IoT Networks 211
Shugang Hao, Lingjie Duan

Harnessing Contrastive Learning and Neural Transformation for Time Series Anomaly Detection..... 216
Jiazhen Chen, Mingbin Feng, Tony S. Wirjanto

Knowledge Distillation for Image Restoration : Simultaneous Learning from Degraded and Clean Images 221
Yongheng Zhang, Danfeng Yan

Hierarchical Similarity Loss Enhanced Depth and Structural Fidelity in Monocular RGB-To-Depth Mapping with Adversarial Training..... 226
Changzeng Fu, Yikai Su, Kaifeng Su, Le Yang, Peng Shan, Xiaoyong Lv, Yuliang Zhao

Global Context MambaVision for EEG-Based Emotion Recognition..... 231
Hao Wang, Li Xu, Yuntao Yu, Weiyue Ding, Yiming Xu

Automated Exposure Mapping for Networked Interference 236
Yunxin Mao, Haotian Wang, Yishuai Cai, Minglong Li, Ji Wang, Wenjing Yang

BiMA: Bidimensional Multi-Level Attention Embedded Network for Single-Frame Infrared Small Target Detection 241
He Deng, Xiaojie Yin, Xianmin Lan

Revelio: A Real-World Screen-Camera Communication System with Visually Imperceptible Data Embedding..... 246
Abbaas Alif Mohamed Nishar, Shrinivas Kudekar, Bernard Kintzing, Ashwin Ashok

Bridging the Fairness Gap: Enhancing Pre-Trained Models with LLM-Generated Sentences 251
Liu Yu, Ludie Guo, Ping Kuang, Fan Zhou

TGDrag: Adding Semantic Control into Point-Based Image Editing Via Text Guidance 256
Chenhao Lin, Yanjie Zhu, Yingmao Miao, Zhengyu Zhao, Shuai Liu, Chao Shen

Active Visual Learning for Robots with Dueling Deep Q-Networks and Transformer Encoders 261
Hongliang Zeng, Ping Zhang, Fang Li, Qinpeng Yi, Jiahua Wang, Tingyu Ye

Look Before You Leap: Problem Elaboration Prompting Improves Mathematical Reasoning in Large Language Models 266
Haoran Liao, Jidong Tian, Shaohua Hu, Zhihao Zhu, Hao He, Yaohui Jin

Efficient Infrared Image Super-Resolution Reconstruction Via Guided Filter Coefficients Estimation with Parallax Attention Mechanism 271
Qingyao Wu, Bosheng Chen, Chen Li, Xiaotong Tu, Xinghao Ding, Yue Huang

HFE-RWKV: High-Frequency Enhanced RWKV Model for Efficient Left Ventricle Segmentation in Pediatric Echocardiograms 276
Zi Ye, Tianxiang Chen, Ziyang Wang, Hanwei Zhang, Lijun Zhang

Practical Radar Sensing Using Two Stage Neural Network for Denoising OTFS Signals 281
Ashok S Kumar, Sheetal Kalyani

Gaussian Constrained Diffeomorphic Deformation Network for Panoramic Semantic Segmentation 286
Jing Jiang, Jiankun Zhu, Zhaopan Xu, Xi Chen, Sicheng Zhao, Hongxun Yao

FedDiT: Federated Learning by Distillation Token Enhanced Vision Transformer 291
Jue Xiao, Zepu Yi, Hewang Nie, Zhi Lu, Xueming Tang, Songfeng Lu, Zhiguo Huang, Runqing Zhang

Knowledge-Guided Prompt Learning for Deepfake Facial Image Detection 296
Hao Wang, Cheng Deng, Zhidong Zhao

Improving Zero-Shot Chinese-English Code-Switching ASR with kNN-CTC and Gated Monolingual Datastores 301
Jiaming Zhou, Shiwan Zhao, Hui Wang, Tian-Hao Zhang, Haoqin Sun, Xuechen Wang, Yong Qin

Unsupervised Hierarchical Dynamic Similarity Hashing for Multimedia Retrieval 306
Yunfei Chen, Zhan Yang, Jun Long

RemoteTrimmer: Adaptive Structural Pruning for Remote Sensing Image Classification 311
Guangwenjie Zou, Liang Yao, Fan Liu, Chuanyi Zhang, Xin Li, Ning Chen, Shengxiang Xu, Jun Zhou

MonTransformer: Self-Supervised Phonetic to Glyph Conversion Leveraging Positional Context for Traditional Mongolian Texts 316
Chenyang Zhou, Monghjaya, Licheng Wu

M2R-Whisper: Multi-Stage and Multi-Scale Retrieval Augmentation for Enhancing Whisper 321
Jiaming Zhou, Shiwan Zhao, Jiabei He, Hui Wang, Wenjia Zeng, Yong Chen, Haoqin Sun, Aobo Kong, Yong Qin

V-Fusion: 2D Detection-Enhanced Multimodal 3D BEV Object Detection 326
Zhenpeng Li, Xiao Zhao, Jingwei Bian, Biao Liu, Wei Li, Lihua Zhang

Less Over More: Interference Sample Gradient Purification for Parallel Continual Learning 331
Tingyang Lu, Jiayao Tan, Linyan Li, Fuyuan Hu

Chat-Driven 3D Human Pose and Shape Editing with Large Language Models 336
Feng Zhou, Chi Li, Ju Dai, Mengxiao Zhu, Yongmei Zhang, Yu-Kun Lai, Paul L. Rosin

A Pre-Trained Plug-In Mixture-Of-LoRAs Model for Transferable Sequential Recommendation 341
Wenqi Sun, Ruobing Xie, Junjie Zhang, Zitian Guo, Wayne Xin Zhao, Zhanhui Kang, Ji-Rong Wen

Segment-Recurrent Transformer with Multi-Scale Fusion for Long-Term Time Series Forecasting 346
Ziang Yang, Lingwei Wei, Biyu Zhou, Xuehai Tang, Ruixuan Li, Songlin Hu

Tip the Scales: Achieving Balance in Adversarial Examples Across Modalities 351
Zhenbo Shi, Zhidong Yu, Yuxuan Zhang, Shuchang Wang, Xiaoman Liu, Wei Yang, Liusheng Huang

Enhancing Remote Sensing Vision-Language Models for Zero-Shot Scene Classification 356
Karim El Khoury, Maxime Zanella, Benoît Gérin, Tiffanie Godelaine, Benoît Macq, Saïd Mahmoudi, Christophe De Vleeschouwer, Ismail Ben Ayed

SUFT: Sparse and Uncertain Fusion Transformers for Multi-Atlas Brain Network Analysis 361
Zhan Su, Jiashuang Huang, Shu Jiang, Mingliang Wang, Weiping Ding

Leveraging Boolean Directivity Embedding for Binaural Target Speaker Extraction 366
Yichi Wang, Jie Zhang, Chengqian Jiang, Weitai Zhang, Zhongyi Ye, Lirong Dai

Integrating Failures in Robot Skill Acquisition with Offline Action-Sequence Diffusion RL 371
Hecheng Wang, Lizhe Qi, Yunquan Sun

Dual-PST: Dual-Branch SpatioTemporal-Planar Network for Video Forgery Detection 376
Siyu Liu, Zhida Zhang, Junxian Duan, Jie Cao, Aihua Zheng

A Robust Online Miscalibration Detection and Correction Method for LiDAR-Camera 381
Feng Pan, Wei Wang, Jianing Zhang

DiffGAP: A Lightweight Diffusion Module in Contrastive Space for Bridging Cross-Model Gap 386
Shentong Mo, Zehua Chen, Fan Bao, Jun Zhu

Integrating Pause Information with Word Embeddings in Language Models for Alzheimer's Disease Detection from Spontaneous Speech 391
Yu Pu, Wei-Qiang Zhang

Consensus Graph Filter Learning for Multiple Graph Clustering 396
Jiale Zou, Yan Chen, Peng Zhou, Chao Wen, Liang Du, Yuhua Qian

Advancing Dark Action Recognition Via Modality Fusion and Dark-To-Light Diffusion Model 401
Yuxuan Wang, Zhen Xing, Zuxuan Wu

Robust Detection Based on the K-Score Test 406
Koby Todros

Training-Free Task Planning by Parsing Language Signals with Common Sense 411
Xianqi Zhang, Wenrui Wang, Shitong Chai, Xingtao Wang, Xiaopeng Fan

Keypoint Aware Masked Image Modelling 416
Madhava Krishna, A. V. Subramanyam

Enhancing Vision: Harmonizing Frequency for Imaging Quality and Perception Accuracy 421
Hongyang Chen, Kaisheng Ma

CPA-Enhancer: Chain-Of-Thought Prompted Adaptive Enhancer for Downstream Vision Tasks Under Unknown Degradations 426
Yuwei Zhang, Yan Wu, Yanming Liu, Xinyue Peng

Gated Cross-Attention Network for Depth Completion 431
Xiaogang Jia, Songlei Jian, Yusong Tan, Yonggang Che, Wei Chen, Zhengfa Liang

Modulo Sampling and Recovery with Unknown and Time-Varying Folding Parameter 436
Yhonatan Kvich, Alperen Yasar, Eyyup Tasci, Rabia T. Yazicigil, Yonina C. Eldar

Synthetic Dataset Generation for String Ensemble Separation 441
Minju Kim, Joonhyeon Bae, Eunsik Shin, Kyogu Lee

PaSTS: Parameter-Affined Seasonal-Trend Synthesis for Multi-Dimensional Long-Term Time Series Forecasting Within LLM 446
Quanfeng Lv, Jingguo Ge, Yifei Xu, Tong Li, Liangxiong Li

Developing a Multilingual Dataset and Evaluation Metrics for Code-Switching: A Focus on Hong Kong's Polylingual Dynamics 451
Peng Xie, Kani Chen

Spiking Transformer with Spatial-Temporal Spiking Self-Attention 456
Zhaokun Zhou, Jun Niu, Yang Zhang, Li Yuan, Yuesheng Zhu

SpikingPoint: Rethinking Point as Spike for Efficient 3D Point Cloud Analysis 461
Zhaokun Zhou, Yijie Lu, Jiaqiyu Zhan, Guibo Luo, Yuesheng Zhu

Vision-Language Model Guided Semi-Supervised Learning for No-Reference Video Quality Assessment 466
Shankhanil Mitra, Rajiv Soundararajan

Segue: Side-Information Guided Generative Unlearnable Examples for Facial Privacy Protection in Real World 471
Zhiling Zhang, Jie Zhang, Kui Zhang, Wenbo Zhou, Ting Xu, Daiheng Gao, Zixian Guo, Qinglang Guo, Weiming Zhang, Nenghai Yu

Part in Part Embedding Network for Zero-Shot Learning 476
Zhexian Zhou, Liang Xiao, Guo-Sen Xie

Federated Smoothing ADMM for Quantile Regression with Non-Convex Sparse Penalties 481
Reza Mirzaeifard, Stefan Werner

Tool Playgrounds: A Comprehensive and Analyzable Benchmark for LLM Tool Invocation 486
Zhiwei Dong, Ruihao Gong, Yang Yong, Shuo Wu, Yongqiang Yao, Song-Lu Chen, Xu-Cheng Yin

Distilling Generative-Discriminative Representations for Very Low-Resolution Face Recognition 491
Junzheng Zhang, Weijia Guo, Bochao Liu, Ruixin Shi, Yong Li, Shiming Ge

Spectral Low-Rank Attention with Flow-Based Refinement for Spectral Reconstruction 496
Yiwen Wang, Zixin Tang, Yexun Hu, Guisong Liu, Tai-Xiang Jiang

Scalable Speech Enhancement with Dynamic Channel Pruning 501
Riccardo Miccini, Clément Laroche, Tobias Piechowiak, Luca Pezzarossa

Joint Space-Time Adaptive Processing and Beamforming Design for Cell-Free ISAC Systems 506
Rang Liu, Ming Li, Qian Liu

Speech Retrieval-Augmented Generation Without Automatic Speech Recognition 511
Do June Min, Karel Mundnich, Andy Lapastora, Erfan Soltanmohammadi, Srikanth Ronanki, Kyu Han

GREST: Ghost Targets Removal Algorithm Using Multipath Angle Estimation 516
Ryuhei Takahashi, Pu Wang

Separate Estimation of Angular Velocity and Angle for Digital Array Radar 521
Tsubasa Terada, Toshihiro Ito, Ryuhei Takahashi

SSDViT: Exploring Siamese and Self Distillation in ViTs for Generalizable Person Re-Identification 526
Jieru Jia, Jianchao Yang

Secure Analog Beamforming Design for Wireless Communication Systems with Movable Antennas 531
Weijie Xiong, Kai Zhong, Zhiling Xiao, Jingran Lin, Qiang Li

BIAWDiff: Enhancing Low-Light Images with Bio-Inspired Attention and Wavelet Diffusion 536
Zeyu Li, Sheng Yang, Hanxiang Yang, Xiongxin Tang, Fengge Wu, Fanjiang Xu

Uncertainty-Participation Context Consistency Learning for Semi-Supervised Semantic Segmentation 541
Jianjian Yin, Yi Chen, Zhichao Zheng, Junsheng Zhou, Yanhui Gu

One-Shot Face Avatar Generation in a Single Forward Pass with Identity Preservation 546
Yingmao Miao, Chenhao Lin, Zhengyu Zhao, Hang Wang, Shuai Liu, Chao Shen, Xiaohong Guan

Conservative Offline Meta-Reinforcement Learning with Task Similarity Measurement 551
Haorui Li, Jiaqi Liang, Linjing Li, Daniel Zeng

Agentic Copyright Watermarking Against Adversarial Evidence Forgery with Purification-Agnostic Curriculum Proxy Learning 556
Erjin Bao, Ching-Chun Chang, Hanrui Wang, Isao Echizen

BP-GPT: Auditory Neural Decoding Using fMRI-Prompted LLM 561
Xiaoyu Chen, Changde Du, Che Liu, Yizhe Wang, Huiguang He

DS-BTIAN: A Novel Deep-Shallow Bidirectional Transformer Interactive Attention Network for Multimodal Emotion Recognition 566
Zengzhao Chen, Chuanxu Zhao, Zhifeng Wang, Chuan Liu, Qiuyu Zheng, Cheng Zou

LiSenNet: Lightweight Sub-Band and Dual-Path Modeling for Real-Time Speech Enhancement 571
Haoyin Yan, Jie Zhang, Cunhang Fan, Yeping Zhou, Peiqi Liu

Enhancing Data-Free Class-Incremental Learning Via Image-Centric Dual Distillation 576
Feifei Fu, Zhiwu Lu

A Hierarchical Compression Technique for 3D Gaussian Splatting Compression 581
He Huang, Wenjie Huang, Qi Yang, Yiling Xu, Zhu Li

MAID: Model Attribution Via Inverse Diffusion 586
Luyu Zhu, Kai Ye, Jiayu Yao, Chenxi Li, Luwen Zhao, Yuxin Cao, Derui Wang, Jie Hao

Fourth-Order Cumulant Based 3-D Near-Field Underdetermined Parameter Estimation with Exact Spatial Propagation Model 591
Longsheng Jin, Hua Chen, Jiaxiong Fang, Wei Liu, Ye Tian, Gang Wang

Can Large Language Models Grasp Event Signals? Exploring Pure Zero-Shot Event-Based Recognition 596
Zongyou Yu, Qiang Qu, Xiaoming Chen, Chen Wang

Contrastive Learning Via Randomly Generated Deep Supervision 601
Shibo Wang, Zili Ma, Ka-Hou Chan, Yue Liu, Tong Tong, Qinquan Gao, Guangtao Zhai, Xiaohong Liu, Tao Tan

Fusion-OSR: Cross-Domain Contrastive Learning with Weibull Calibration for Time Series Open Set Recognition 606
Shuguo Hu, Xudong Zhao, Shuwei Hu, Xuan Gao

Learning Markup Language Model for Composite Relationships Extraction 611
Fengyu Lu, Jiaxin Duan, Junfei Liu

A Weakly Supervised Semantic Segmentation Model with Enhanced CLIP Feature Extraction 616
Fanxuan Kong, Jun Lu

OCTAMamba: A State-Space Model Approach for Precision OCTA Vasculature Segmentation 621
Shun Zou, Zhuo Zhang, Guangwei Gao

BDCKD: Unlocking the Power of Brownian Distance Covariance in Knowledge Distillation 626
Guoming Lu, Heng Yin, Zhiyong Shu, Jielei Wang, Guangchun Luo

VibeGait: Enhancing Structural-Vibration Based Gait Recognition Using Vision 631
Mainak Chakraborty, Chandan, Bodhibrata Mukhopadhyay, Sahil Anchal, Subrat Kar

Diverse Collaboration in Multi-Agent Reinforcement Learning Via Self-Adaptive Method 636
Xiang Xue, Quan Liu, Meilong Shi, Yuchao Jin

SwapTalk: Audio-Driven Talking Face Generation with One-Shot Customization in Latent Space 641
Zeren Zhang, Haibo Qin, Jiayu Huang, Jo-Ku Cheng, Yixin Li, Hui Lin, Yitao Duan, Jinwen Ma

MPAM-3DGS: Multi-Parametric Adversarial Manipulation for 3D Gaussian Splatting 646
Wenxiang Jiang, Hanwei Zhang, Weigang Wang, Zhongwen Guo, Tianao Zhang, Hao Wang

Deep Unfolding of Full Waveform Inversion for Quantitative Ultrasound Imaging 651
Niv Cohen, Yhonatan Kvich, Rui Guo, Yonina C. Eldar

BS-Breath: Respiration Sensing with Cell-Free Massive MIMO 656
Haoqiu Xiong, Robbert Beerten, Zhuangzhuang Cui, Yang Miao, Sofie Pollin

Subspace-Based Range-Angle Tracking for Coherent FDA Radar 661
Yan Sun, Wen-Qin Wang, Maria Sabrina Greco, Fulvio Gini

Causal fMRI-Mamba: Causal State Space Model for Neural Decoding and Brain Task States Recognition 666
Weihao Deng, Fei Han, Qinghua Ling, Qing Liu, Henry Han

Generating Editable Head Avatars with 3D Gaussian GANs 671
Guohao Li, Hongyu Yang, Yifang Men, Di Huang, Weixin Li, Ruijie Yang, Yunhong Wang

Towards Interactive Deepfake Analysis 676
Lixiong Qin, Ning Jiang, Yang Zhang, Yuhan Qiu, Dingheng Zeng, Jiani Hu, Weihong Deng

Detecting and Defending Against Adversarial Attacks on Automatic Speech Recognition Via Diffusion Models 681
Nikolai L. Kühne, Astrid H. F. Kitchena, Marie S. Jensen, Mikkel S. L. Brøndt, Martin Gonzalez, Christophe Biscio, Zheng-Hua Tan

Controllable Forgetting Mechanism for Few-Shot Class-Incremental Learning 686
Kirill Paramonov, Mete Ozay, Eunju Yang, Jijoong Moon, Umberto Michieli

Advancing Active Speaker Detection for Egocentric Videos 691
Jaesung Huh, Juan Azcarreta Ortiz, Anurag Kumar, Ashutosh Pandey, Ali Aroudi, Daniel D.E. Wong, Francesco Nesta, Buye Xu, Jacob Donley

Rethinking the Fragility and Robustness of Fingerprints of Deep Neural Networks 696
Fangqi Li, Shilin Wang, Lei Yang

Emotion-Aware Structural Enhancement Graph Auto-Encoder for Rumor Detection 701
Guoyi Li, Zhongjiang Yao, Die Hu, Yingrui Xu, Xiaodan Zhang, Honglei Lyu

Synergistic Spotting and Recognition of Micro-Expression Via Temporal State Transition 706
Bochao Zou, Zizheng Guo, Wenfeng Qin, Xin Li, Kangsheng Wang, Huimin Ma

Interpreting Deep Neural Network-Based Receiver Under Varying Signal-To-Noise Ratios 711
Marko Tuononen, Dani Korpi, Ville Hautamäki

Multi-Modal Medical Image Fusion Via 3D Manifold Fitting and Dual-Domain Cross-Attention 716
Zeyu Wang, Jiayu Wang, Haiyu Song, Jiawei Feng, Haoran Duan

A Frequency-Aware Augmentation Network for Mental Disorders Assessment from Audio 721
Shuanglin Li, Siyang Song, Rajesh Nair, Syed Mohsen Naqvi

Bootstrapping Language-Audio Pre-Training for Music Captioning 726
Luca A. Lanzendörfer, Constantin Pinkl, Nathanaël Perraudin, Roger Wattenhofer

Speaking Without Sound: Multi-Speaker Silent Speech Voicing with Facial Inputs Only 731
Jaejun Lee, Yoori Oh, Kyogu Lee

The Sound of Water: Inferring Physical Properties from Pouring Liquids 736
Piyush Bagad, Makarand Tapaswi, Cees G. M. Snoek, Andrew Zisserman

Realistic Real-Time Talking Head Synthesis with Grid Encoding and Progressive Conditioning 741
Zhiling Ye, Liang-Guo Zhang, Dingheng Zeng, Quan Lu, Ning Jiang

AKI360: Enabling Highly Interactive 360-Degree Video Streaming by Adaptive Keyframe Interval 746
Haitao Liu, Xinyi Zhang, Chuanmin Jia, Yanbiao Li, Gaogang Xie

Nonlinear Anisotropic Diffusion-Based Channel Estimation in 5G Wireless Networks 751
Kanwardeep Singh Gahlot, Sandeep Joshi, Ke Wang

MorphFader: Enabling Fine-Grained Controllable Morphing with Text-To-Audio Models 756
Purnima Kamath, Chitralekha Gupta, Suranga Nanayakkara

LipGen: Viseme-Guided Lip Video Generation for Enhancing Visual Speech Recognition 761
Bowen Hao, Dongliang Zhou, Xiaojie Li, Xingyu Zhang, Liang Xie, Jianlong Wu, Erwei Yin

Granularity-Aware Contrastive Learning for Fine-Grained Action Recognition 766
Hailun Zhang, Xinrui Wang, Qijun Zhao

AD2T: Adversarial Distortion Domain Translation for Robust Watermarking Against Non-Differentiable Distortions 771
Chengxin Zhao, Hefei Ling, Jiazhong Chen, Han Fang, Zongyi Li, Sijing Xie

Adversarial Feature Disentanglement Framework for Voice Pathology Detection 776
Yifan Xiong, Dongyue Guo, Lipeng Shen, Wei Mo, Hui Yang, Yi Lin

Enhancing Extrapolation Reasoning on Temporal Knowledge Graphs with Logic Rules and Queries 781
Tingxuan Chen, Liu Yang, Zidong Wang, Shuai Luo, Jun Long

Dual-Pyramid Attention Collaborative Network for Oracle Bone Inscription Classification 786
Jiaying Gao, Fausto Giunchiglia, Tongyu Zhao, Chuntao Li, Hao Xu

Multi-Prototype-Based Embedding Refinement for Medical Image Segmentation 791
Yali Bi, Enyu Che, Yinan Chen, Yuanpeng He, Jingwei Qu

LSU-NET: Lightweight Automatic Organs Segmentation Network for Medical Images 796
Yujie Ding, Shenghua Teng, Zuoyong Li, Xiao Chen

Zero-Shot Document Retrieval with Hybrid Pseudo-Document Retriever 801
Dong Sun, Wenya Guo, Xumeng Liu, Ying Zhang, Zhaoxiang Hou, Zengxiang Li

Electrocardiogram Report Generation and Question Answering Via Retrieval-Augmented Self-Supervised Modeling 806
Jialu Tang, Tong Xia, Yuan Lu, Cecilia Mascolo, Aaqib Saeed

Robust Activity Detection for Massive Access Using Covariance-Based Matching Pursuit 811
Xinjue Wang, Esa Ollila, Sergiy A. Vorobyov

A Key to Effective Multi-Task Learning: Separate Query Selection for Task-Synergized Handling and Node Utilization 816
Shan-Ya Yang, Hao-Chung Cheng, Chien-Yao Wang, Jia-Ching Wang, Chun-Yi Lee

Micro-Expression Spotting Based on Multi-Modal Hierarchical Semantic-Guided Deep Fusion Model* 821
Zhihua Xie, Haolin Chang

Discriminating Mizo Hunting and War Chants Using Acoustic Features 826
Esther Ramdinmawii, Vinay Kumar Mittal

Content and Salient Semantics Collaboration for Cloth-Changing Person Re-Identification 831
Qizao Wang, Xuelin Qian, Bin Li, Lifeng Chen, Yanwei Fu, Xiangyang Xue

GS-PT: Exploiting 3D Gaussian Splatting for Comprehensive Point Cloud Understanding Via Self-Supervised Learning 836
Keyi Liu, Yeqi Luo, Weidong Yang, Jingyi Xu, Zhijun Li, Wen-Ming Chen, Ben Fei

Out-Of-Distribution Detectors: Not Yet Primed for Practical Deployment 841
Changshun Wu, Wendi Ding, Xiaowei Huang, Saddek Bensalem

Everyone-Can-Sing: Zero-Shot Singing Voice Synthesis and Conversion with Speech Reference 846
Shuqi Dai, Yunyun Wang, Roger B. Dannenberg, Zeyu Jin

Instance–wise Feature Acquisition with Classifier Selection Option for Structured Data Instances 851
Sachini Piyoni Ekanayake, Daphney–Stavroula Zois

Cross-Component Residual Prediction for Geometry-Based Point Cloud Compression 856
Bharath Vishwanath, Yingzhan Xu, Kai Zhang, Li Zhang

Selective Attention Merging for Low Resource Tasks: A Case Study of Child ASR 861
Natarajan Balaji Shankar, Zilai Wang, Eray Eren, Abeer Alwan

Interference-Resilient Hybrid Multi-Antenna ARQ 866
Nicholas D. Sidiropoulos, Yuanyuan Tang

Physics-Informed Neural Networks for Ocean Acoustic Field Prediction with Envelope Smoothing 871
Yongsung Park, Peter Gerstoft, Seunghyun Yoon, Woojae Seong

Sharpness-Aware Minimization with Adaptive Regularization for Training Deep Neural Networks 876
Jinping Zou, Xiaoge Deng, Tao Sun

Lead Instrument Detection from Multitrack Music 881
Longshen Ou, Yu Takahashi, Ye Wang

Debiased Training for Semi-Supervised Sound Event Detection 886
Shengchang Xiao, Xueshuai Zhang, Pengyuan Zhang, Yonghong Yan

Hierarchical Proxy Learning for Cloth-Changing Person Re-Identification 891
Chenyang Yu, Xuehu Liu, Ju Dai, Pingping Zhang, Huchuan Lu

A Cost-Effective Solution for Remote Sensing Image Segmentation Via Train/Test-Time Adaptation 896
Wei Chen, Xin Luo, Yulin He, Tianrui Liu, Di Wu, Tianhang Guo, Yuhang Li, Yuhua Tang

Fooling the Forgers: A Multi-Stage Framework for Audio Deepfake Detection 901
Gautam Siddharth Kashyap, Zohaib Hasan Siddiqui, Mohammad Anas Azeez, Rafiq Ali, Shantanu Kumar, Navin Kamuni, Jiechao Gao

Breaking Through the Spike: Spike Window Decoding for Accelerated and Precise Automatic Speech Recognition 906
Wei Zhang, Tian-Hao Zhang, Chao Luo, Hui Zhou, Chao Yang, Xinyuan Qian, Xu-Cheng Yin

Graph Structure Learning Via Transfer Entropy for Multivariate Time Series Anomaly Detection 911
Mingyu Liu, Yijie Wang, Xiaohui Zhou, Yongjun Wang

Can AI See What We Can't? Leveraging Deep Learning and Multi-Temporal Satellite Data to Revolutionize Crop Type Mapping and Yield Prediction 916
Gautam Siddharth Kashyap, Harsh Joshi, Manaswi Kulahara, Rajkumar Dhakar, Atul Sajjanhar, Jiechao Gao, Sarthak Jain, Shahab Saquib Sohail

Exploring Temporal Constraints for Unsupervised Iris Motion Tracking in AS-OCT Videos 921
Lingxi Hu, Xiao Wu, Risa Higashita, Xiaoli Xing, Menglan Zhou, Song Lin, Xiaorong Li, Xiaoling Li, Jinming Duan, Jiang Liu

Local Statistics for Generative Image Detection 926
Yung Jer Wong, Teck Khim Ng

Modality Modulation and Dual Consistency for Multi-Modality Semi-Supervised Medical Image Segmentation 931
Yingyu Chen, Ziyuan Yang, Deng Xiong, Yi Zhang

Enhancing Emotion Reasoning for Image Multi-Emotion Prediction 936
Bingbing Wang, Geng Tu, Bin Liang, Zhixin Bai, Min Yang, Xi Zeng, Liang Yao, Ruifeng Xu

Multi-View Feature Discrepancy Attack for Single Object Tracking 941
Zhiheng Li, Weng Zhimin, Yuehuan Wang

TopoRefine: Iterative Refinement with Reasoning Topology as High-Level Feedback 946
Haoran Liao, Shaohua Hu, Zhihao Zhu, Hao He, Yaohui Jin

Retinex-Based Self-Conditioned Diffusion Model for Low-Light Image Enhancement 951
Jiawei Zhang, Ziwen Li, Jinpu Zhang, Yuehuan Wang

Faithful Self-Refinement in Mathematical Reasoning Via Progressive Back-Translation 956
Haoran Liao, Zhihao Zhu, Shaohua Hu, Hao He, Yaohui Jin

Map-Guided Few-Shot Audio-Visual Acoustics Modeling 961
Diwei Huang, Kunyang Lin, Peihao Chen, Qing Du

Unveiling Local Well-Posedness Influence for Cross-Modal Person Re-Identification 966
Yumeng Yang, Guan-Nan Dong, Aichun Zhu, Mingcheng Ni, Yifeng Li

G-Depth: An Efficient Graph Method for Robust Depth Completion 971
Zhongyu Huang, Yijun Chen, Aizierjiang Aiersilan, Li Li

Complementary Graph Learning and Prompt-Based Cross-Modal Generation for Missing-Modality Fake News Detection 976
Fei Wu, Ruixuan Zhou, Changhui Hu, Qinghua Huang, Xiao-Yuan Jing

ComplexDec: A Domain-Robust High-Fidelity Neural Audio Codec with Complex Spectrum Modeling 981
Yi-Chiao Wu, Dejan Marković, Steven Krenn, Israel D. Gebru, Alexander Richard

Explainable Detection of Alzheimer's Disease Through Analysis of Human Behavior in Video 986
Bao-Hsuan Huang, Po-Chih Kuo, Likai Huang, Chaur-Jong Hu, Cheng-Yu Chen

A Robust Distributed Recurrent Neural Network for Multi-Agent Consensus Control 991
Yiwei Li, Jiaxin Liu, Lin Yang, Yating Zhang, Kunlin Liu, Ge Zhou, Liangze Yin, Wei Dong

Signal Processing Challenges in Automotive Radar 996
Sandeep Rao, Rajan Narasimha, Shunqiao Sun

STA-V2A: Video-To-Audio Generation with Semantic and Temporal Alignment 1001
Yong Ren, Chenxing Li, Manjie Xu, Wei Liang, Yu Gu, Rilin Chen, Dong Yu

Improving Acoustic Scene Classification in Low-Resource Conditions 1006
Zhi Chen, Yun-Fei Shao, Yong Ma, Mingsheng Wei, Le Zhang, Wei-Qiang Zhang

Spectral Enhancement and Pseudo-Anchor Guidance for Infrared-Visible Person Re-Identification 1011
Yiyuan Ge, Zhihao Chen, Ziyang Wang, Jiaju Kang, Mingya Zhang

WMRE: Enhancing Distant Supervised Relation Extraction with Word-Level Multi-Instance Learning and Multi-Hierarchical Feature 1016
Xiang Ying, Xiangchuan Xie, Tianyi Xu, Yue Zhao, Zechen Meng, Mankun Zhao

SkeletonMix: A Mixup-Based Data Augmentation Framework for Skeleton-Based Action Recognition 1021
Zongye Zhang, Huanyu Zhou, Qingjie Liu, Yunhong Wang

Exploring the Interpretability of EEG-Inception Convolutional Neural Networks for Epilepsy Prediction 1026
Guanglong Zhang, Tianren Wang, Jinjie Guo, Zhiyuan Yang, YiLian Wu, Guixia Kang

Self-Optimization Training for Weakly Supervised Image Manipulation Localization 1031
Zhangchen Zhu, Jiafeng Li, Ying Wen

DCD-MUSIC: Deep-Learning-Aided Cascaded Differentiable MUSIC Algorithm for Near-Field Localization of Multiple Sources 1036
Arad Gast, Luc Le Magoarou, Nir Shlezinger

A Reinforcement Learning Agent Controlled Multi-Branch Small Object Detection Framework..... 1041
Junkun Hong, Yitian Long, Yueyi Luo, Liujie Hua, Jun Long, Qianqian Qi

HieClip: Hierarchical CLIP with Explicit Alignment for Zero-Shot Anomaly Detection 1046
Liujie Hua, Xiu Su, Yueyi Luo, Shan You, Jun Long

UME: Upcycling Mixture-Of-Experts for Scalable and Efficient Automatic Speech Recognition..... 1051
Li Fu, Shanyong Yu, Siqi Li, Lu Fan, Youzheng Wu, Xiaodong He

Extending MPR for Locating a Moving Object Based on TDOA and FDOA..... 1056
Beichuan Tang, Yimao Sun, Xiantao Heng, Yanbing Yang, Liangyin Chen

Seek and Solve Reasoning for Table Question Answering..... 1061
Ruya Jiang, Chun Wang, Weihong Deng

TS-Net: Assembling Task-Specific Features from Multiple Feature Levels for Multi-Task Learning..... 1066
Chen Liu, Zhaolin Wan, Penghong Wang, Xingtao Wang, Xiaopeng Fan

Hybrid Content Caching Empowered by AIGC in Wireless Networks 1071
Ding Xu, Lingjie Duan, Hongbo Zhu

Reduced Effectiveness of Kolmogorov-Arnold Networks on Functions with Noise 1076
Haoran Shen, Chen Zeng, Jiahui Wang, Qiao Wang

Learned Approximated Optimization for Rapid Low-Complexity Hybrid Beamforming Design 1081
Amit Milstein, Tomer Yablonka, Nir Shlezinger

ECG-Guided Individual Identification Via PPG..... 1086
Riling Wei, Hanjie Chen, Kelu Yao, Chuanguang Yang, Jun Wang, Chao Li

Atom-Constrained Maximum Likelihood Gridless DOA with Wirtinger Gradients 1091
Peter Gerstoft, Yongsung Park

Developing Instruction-Following Speech Language Model Without Speech Instruction-Tuning Data 1096
Ke-Han Lu, Zhehuai Chen, Szu-Wei Fu, Chao-Han Huck Yang, Jagadeesh Balam, Boris Ginsburg, Yu-Chiang Frank Wang, Hung-Yi Lee

MoWE-Audio: Multitask AudioLLMs with Mixture of Weak Encoders 1101
Wenyu Zhang, Shuo Sun, Bin Wang, Xunlong Zou, Zhuohan Liu, Yingxu He, Geyu Lin, Nancy F. Chen, Ai Ti Aw

DRCap: Decoding CLAP Latents with Retrieval-Augmented Generation for Zero-Shot Audio Captioning 1106
Xiquan Li, Wenxi Chen, Ziyang Ma, Xuenan Xu, Yuzhe Liang, Zhisheng Zheng, Qiuqiang Kong, Xie Chen

Identity-Agnostic Learning for Deepfake Face Detection 1111
Xuan Zhou, Zongyong Deng, Qijun Zhao

VN-GT: Optimizing Virtual Network Deployment Via Game Theory 1116
Weijie Wang, Yan Wang, Guokun Xu, Zuxin Chen, Siyuan Li, Min Yu, Weiqing Huang, Degang Sun

MUPO-Net: A Multilevel Dual-Domain Progressive Enhancement Network with Embedded Attention for CT Metal Artifact Reduction 1121
Xiaoli Yao, Jia Tan, Zijian Deng, Deng Xiong, Qijun Zhao, Min Wu

Efficient Hierarchical Domain Adaptive Thermal Infrared Tracking 1126
Qiao Li, Kanlun Tan, Qiao Liu, Di Yuan, Xin Li, Yunpeng Liu

DASSL: Domain Agnostic Self-Supervised Learning with Multiple Missing Information Reconstruction Branches 1131
Jiang Fang, Haonan He, Chen Guo, Jiyan Sun, Zhaorui Guo, Chao Xu, Mohan Su, Yinlong Liu, Wei Ma

Harnessing Dimensional Contrast and Information Compensation for Sentence Embedding Enhancement 1136
Kang He, Yuzhe Ding, Bobo Li, Haining Wang, Fei Li, Chong Teng, Donghong Ji

Dual Encoders for Diffusion-Based Image Inpainting 1141
Dezhi Zheng, Kaijun Deng, Jinbao Wang, Linlin Shen

Hypergraph-Based Dynamic Graph Node Classification 1146
Xiaoxu Ma, Chen Zhao, Minglai Shao, Yujie Lin

Bidirectional Reference Image Quality Assessment Via Content-Quality Correlation Modeling 1151
Bo Hu, Wenzhi Chen, Chunyi Li, Jiaxu Leng, Weisheng Li, Xinbo Gao

PDSeg: Patch-Wise Distillation and Controllable Image Generation for Weakly-Supervised Histopathology Tissue Segmentation 1156
Wei-Hua Li, Yu-Hsing Hsieh, Huei-Fang Yang, Chu-Song Chen

Redefining Well Exposedness for Locally Adaptive Multi-Exposure Fusion 1161
Prince Arya, Saurabh Kumar, Ashish Agarwal, Nutan Yenneti, Narasimha Pai

GEGA: Graph Convolutional Networks and Evidence Retrieval Guided Attention for Enhanced Document-Level Relation Extraction 1166
Yanxu Mao, Xiaohui Chen, Peipei Liu, Tiehan Cui, Zuhui Yue, Zheng Li

Learn from Balance: Rectifying Knowledge Transfer for Long-Tailed Scenarios 1171
Xinlei Huang, Jialiang Tang, Xubin Zheng, Jinjia Zhou, Wenxin Yu, Ning Jiang

Low-Complexity Own Voice Reconstruction for Hearables with an In-Ear Microphone 1176
Mattes Ohlenbusch, Christian Rollwage, Simon Doclo

Global Tropical Cyclone Intensity Forecasting with Multi-Modal Multi-Scale Causal Autoregressive Model 1181
Xinyu Wang, Kang Chen, Lei Liu, Tao Han, Bin Li, Lei Bai

An Adaptive Framework for Multi-View Clustering Leveraging Conditional Entropy Optimization 1186
Lijian Li, Yuanpeng He, Chi-Man Pun

Multi-Label Body Constitution Recognition Via Dual Transform MLP-Like Architecture Using Tongue Images....1191
Mengjian Zhang, Guihua Wen, Pei Yang

Graph-Driven Insights: Enhancing Stock Market Prediction with Relational Temporal Dynamics....1196
Renjun Jia, Kaiming Yang, Dawei Cheng, Li Han, Yuqi Liang

Training Better Embedding with Perturbed Data Augmentation for Automatic Singing Quality Assessment 1201
Po-Wei Chen, Von-Wun Soo

RetinaStereo: Dynamic-Volume Stereo Matching Network 1206
Xiaoyan Liao, Haoliang Zhao, Fan Yang, Kwokching Cheung, Jun Jiang, Yong Zhao, Jie Chen, Xinan Wang

One-Step Incomplete Multi-View Clustering Based on Bipartite Graph Learning 1211
Minghao Li, Hechuan Lin, Huiying Xu, Ziying Wang, Xinzhong Zhu, Xiao Huang

Loudspeaker Beamforming to Enhance Speech Recognition Performance of Voice Driven Applications.... 1216
D. de Groot, B. Karslioglu, O. Scharenborg, J. Martinez

Deep Feedback Cancellation for Hearing Aids with Improved System Stability and Sound Quality 1221
Eleftheria Lydaki, Zheng-Hua Tan, Jesper Jensen, Meng Guo

Towards Efficient Deep Hashing Retrieval: Condensing Your Data Via Feature-Embedding Matching.... 1226
Tao Feng, Jie Zhang, Huashan Liu, Zhijie Wang, Shengyuan Pang

Latent Watermarking of Audio Generative Models.... 1231
Robin San Roman, Pierre Fernandez, Antoine Deleforge, Yossi Adi, Romain Serizel

Volatile MAB-Based Configuration Selection for Offloading Video Analytics Tasks to Edges 1236
Yu Liang, Sheng Zhang, Jie Wu

Distributed IRSs Mitigate Spatial Wideband & Beam Split Effects.... 1241
L. Yashvanth, Chandra R. Murthy, Bhaskar D. Rao

Towards Feature-Consistent Parameter Collaboration for Personalized Federated Learning 1246
Xintong Lu, Jiahe Li, Yuchao Zhang, Wendong Wang

Try Before You Buy: Solving Multi-Model Complex Tasks by Model Competitions 1251
Yongqiang Zhao, Zhenyu Li, Zhi Jin, Feng Zhang, Lianwei Wu, Xinhai Xu, Donghong Liu

Enhanced Corneal Endothelial Cell Segmentation Via Frequency-Selected Residual Fourier Diffusion Models.... 1256
Tianyang Wang, Xiaofei Nan, Yunze Wang, Yuhang Yan, Zhenkai Gao, Jingxin Liu

Spatially-Aware Cross-Modal Contrastive Learning for Low-Shot HSI Classification 1261
Akhil Vasim, Pankhi Kashyap, Shabnam Choudhury, Biplab Banerjee

Sampling Nonsmooth Log-Concave Densities: A Comparative Study of Primal-Dual Based Proposal Distributions 1266
Juliette Chevallier, Gersende Fort

A Spectrum-Enhanced Attention Model for Semantic Segmentation of Remote Sensing Images 1271
Xin Li, Feng Xu, Feifei Tao, Yao Tong, Xin Lyu, Jianyi Zhong, André Kaup

When Does Visual Prompting Outperform Linear Probing for Vision-Language Models? a Likelihood Perspective 1276
Hsi-Ai Tsao, Lei Hsiung, Pin-Yu Chen, Tsung-Yi Ho

Accelerating Convergence in Bounding Box Regression with a Refined IoU Loss Function 1281
Enhui Chai, Xingyu Li, Tianxiang Cui, Zheng Lu, Fiseha Berhanu Tesema

Dual Trajectory Revised Diffusion Model for Time Series Forecasting 1286
Zilong Hu, Yan Qiao, Zidang Cai, Rongyao Hu, Junjie Wang, Meng Li, Zhenchun Wei

Ambisonics Binaural Rendering Via Masked Magnitude Least Squares 1291
Or Berebi, Fabian Brinkmann, Stefan Weinzierl, Boaz Rafaely

Rethinking Dual-Stream Super-Resolution for Enhancing Remote Sensing Object Detection 1296
An Luo, Kai Hu, Kai Jiang

Large Language Model Can Transcribe Speech in Multi-Talker Scenarios with Versatile Instructions 1301
Lingwei Meng, Shujie Hu, Jiawen Kang, Zhaoqing Li, Yuejiao Wang, Wenxuan Wu, Xixin Wu, Xunying Liu, Helen Meng

SSFMamba: Spatial-Spectral Fusion State Space Model for Pansharpening 1306
Mengting Ma, Mengjiao Zhao, Yizhen Jiang, Xiangdong Li, Wei Zhang

MSA-ITEI: A Novel Method for Multimodal Analysis of Social Media Stickers 1311
Yuanchen Shi, Fang Kong

Disentangling Speakers in Multi-Talker Speech Recognition with Speaker-Aware CTC 1316
Jiawen Kang, Lingwei Meng, Mingyu Cui, Yuejiao Wang, Xixin Wu, Xunying Liu, Helen Meng

Efficient Streaming LLM for Speech Recognition 1321
Junteng Jia, Gil Keren, Wei Zhou, Egor Lakomkin, Xiaohui Zhang, Chunyang Wu, Frank Seide, Jay Mahadeokar, Ozlem Kalinli

Efficient Co-Approximate Parallel Compressive Depth Reconstruction on FPGA 1326
Yun Wu, John McAllister

Modular Prompt Learning Improves Vision-Language Models 1331
Zhenhan Huang, Tejaswini Pedapati, Pin-Yu Chen, Jianxi Gao

First-Order State Space Model for Lightweight Image Super-Resolution 1336
Yujie Zhu, Xinyi Zhang, Yekai Lu, Guang Yang, Faming Fang, Guixu Zhang

Leveraging Registers in Vision Transformers for Robust Adaptation 1341
Srikar Yellapragada, Kowshik Thopalli, Vivek Narayanaswamy, Wesam Sakla, Yang Liu, Yamen Mubarka, Dimitris Samaras, Jayaraman J. Thiagarajan

Segment Any Bone in CT with Partial Supervision 1346
Tianyou Liang, Xiaoxu Li, Yu Peng, Min Xu

Whitening Effects for ML-DoA Estimation Using a Sparse Representation of Array Covariance 1351
Thomas Aussaguès, Anne Ferréol, Alice Delmer, Pascal Larzabal

Towards Fully Test-Time Adaptation Via Variance Balancing and Semantic Augmentation 1356
Houcheng Su, Bingli Wang, Daixian Liu, Jiao Li, Chen-Bin Feng, Chi-Man Vong

Enhancing Multimodal Emotion Recognition Through Multi-Granularity Cross-Modal Alignment..... 1361
Xuechen Wang, Shiwan Zhao, Haoqin Sun, Hui Wang, Jiaming Zhou, Yong Qin

SpeechCaps: Advancing Instruction-Based Universal Speech Models with Multi-Talker Speaking Style Captioning..... 1366
Chien-Yu Huang, Min-Han Shih, Ke-Han Lu, Chi-Yuan Hsiao, Hung-Yi Lee

FasterGold-DETR: An Efficient End-To-End Fire Detection Model Via Gather-And-Distribute Mechanism..... 1371
Chengming Liu, Fan Wu, Lei Shi

Sparse-VQ Transformer: An FFN-Free Framework with Vector Quantization for Enhanced Time Series..... 1376
Yanjun Zhao, Tian Zhou, Chao Chen, Liang Sun, Yi Qian, Rong Jin

DiffAttack: Diffusion-Based Timbre-Reserved Adversarial Attack in Speaker Identification..... 1381
Qing Wang, Jixun Yao, Zhaokai Sun, Pengcheng Guo, Lei Xie, John H.L. Hansen

SlotFusion: Object-Centric Audiovisual Feature Fusion with Slot Attention for Remote Sensing Scene Recognition..... 1386
Fangzhou Han, Tianyi Yu, Lamei Zhang, Lingyu Si, Yiqi Zhang

Adversarial Knowledge Transfer for Black-Box Model Inversion Attack..... 1391
Xinhao Liu, Zetao Lin, Yingzhao Jiang, Qiao Yan

Efficient Object Placement Via LLM and Diffusion Model..... 1396
Wei Liu, Liuan Wang, Jun Sun

Audio Array-Based 3D UAV Trajectory Estimation with LiDAR Pseudo-Labeling..... 1401
Allen H.-X. Lei, Tianchen Deng, Han Wang, Jianfei Yang, Shenghai Yuan

A Novel Network for Short-Term Wind Speed Prediction: Mitigating Distribution Shift and Feature Loss..... 1406
Mei Yu, Shengkang Dong, Xuewei Li, Zewen Shang, Yingzhou Sun, Zhiqiang Liu

Indoor Sensing with Measurements..... 1411
Vijaya Yajnanarayana, Philipp Geuer, Satyam Dwivedi

TAME: Temporal Audio-Based Mamba for Enhanced Drone Trajectory Estimation and Classification..... 1416
Zhenyuan Xiao, Huanran Hu, Guili Xu, Junwei He

Pruning for Sparse Diffusion Models Based on Gradient Flow..... 1421
Ben Wan, Tianyi Zheng, Zhaoyu Chen, Yuxiao Wang, Jia Wang

UniFaceGAN: High-Quality 3D Face Editing with a Unified Latent Space..... 1426
Jinfu Wei, Zheng Zhang, Ran Liao, Duan Gao

Diffusion Augmentation Sub-Center Modeling for Unsupervised Anomalous Sound Detection with Partially Attribute-Unavailable Conditions..... 1431
Jiawei Yin, Yu Gao, Wenbin Zhang, Tianyi Wang, Mingjun Zhang

Speech Enhancement Using Continuous Embeddings of Neural Audio Codec..... 1436
Haoyang Li, Jia Qi Yip, Tianyu Fan, Eng Siong Chng

LLM-Guided Dual-Branch Diffusion Model for Fine-Grained Motion Synthesis..... 1441
Wenlong Wang, Dahua Gao, Xinyu Liu

Feature Refinement Decomposition and Relation Preference Enhancement for Remote Sensing Change Detection 1446
Wenqi Zheng, Jianing Chen, Junze Yang, Chuhao Chen, Wei Li, Rahul Yadav, Xiangxu Meng

Pre-Training with Synthetic Patterns for Audio..... 1451
Yuchi Ishikawa, Tatsuya Komatsu, Yoshimitsu Aoki

Partial Reconstruction Error for Deepfake Detection..... 1456
Yufei Zhang, Zheling Meng, Bo Peng, Jing Dong, Beilin Chu, Wei Wang

Leveraging Out-Of-Domain Noise for Unsupervised Domain Adaptation in Speech Enhancement 1461
Yu Liao, Haixin Guan, Shuang Wei, Yanhua Long

Birds of a Feather: Learning to Retrieve Dance Poses from Music Via Ground-Truth Annotation Lifting..... 1466
Bo-Wei Tseng, Wen-Li Wei, Jen-Chun Lin

SimulTron: On-Device Simultaneous Speech to Speech Translation..... 1471
Alex Agranovich, Eliya Nachmani, Oleg Rybakov, Yifan Ding, Ye Jia, Nadav Bar, Heiga Zen, Michelle Tadmor Ramanovich

Image Compressive Sensing with Adaptive Sampling by Median Filtering 1476
Yanfeng Wu, Chen Hui, Ronghua Liao, Shaohui Liu, Debin Zhao

Exploring Kolmogorov-Arnold Networks for Realistic Image Sharpness Assessment..... 1481
Shaode Yu, Ze Chen, Zhimu Yang, Jiacheng Gu, Bizu Feng, Qiurui Sun

Improving 5G Positioning Through Signal-To-Noise Ratio Recognition Training..... 1486
Wenqi Zheng, Jianing Chen, Junze Yang, Chuhao Chen, Wei Li, Rahul Yadav, Xiangxu Meng

An Intra- And Cross-Frame Topological Consistency Scheme for Semi-Supervised Atherosclerotic Coronary Plaque Segmentation 1491
Ziheng Zhang, Zihan Li, Dandan Shan, Yuehui Qiu, Qingqi Hong, Qingqiang Wu

PQNAS: Mixed-Precision Quantization-Aware Neural Architecture Search with Pseudo Quantizer..... 1496
Tianxiao Gao, Li Guo, Shihao Wang, Shiai Zhu, Dajiang Zhou

Enhancing Visual Forced Alignment with Local Context-Aware Feature Extraction and Multi-Task Learning 1501
Yi He, Lei Yang, Shilin Wang

Regret Optimization Experience Replay in Off-Policy Reinforcement Learning 1506
Jie Zhang, Yirong Yao, Wei He, Yiqun Niu, Chongjun Wang

PIN: A Prompt-Based Implicit Sentiment Analysis Network for Chinese.....1511
Kun Bu, Yuanchao Liu, Wenbo Wang, Ziyi Cao

MacST: Multi-Accent Speech Synthesis Via Text Transliteration for Accent Conversion..... 1516
Sho Inoue, Shuai Wang, Wanxing Wang, Pengcheng Zhu, Mengxiao Bi, Haizhou Li

EGAS: Enhanced Geometry-Aware 3D Asset Generation Using Gaussian Splatting..... 1521
Shengjie Hu, Xiaogang Zhang, Hua Chen, Wenbin Yan

Nonlinear Precoding in the RIS-Aided MIMO Broadcast Channel 1526
Dominik Semmler, Michael Joham, Wolfgang Utschick

The Importance of Spatial and Spectral Information in Multiple Speaker Tracking 1531
Hanan Beit-On, Vladimir Tourbabin, Boaz Rafaely

Phoneme-Level Contrastive Learning for User-Defined Keyword Spotting with Flexible Enrollment 1536
Kewei Li, Hengshun Zhou, Kai Shen, Yusheng Dai, Jun Du

LiRCDepth: Lightweight Radar-Camera Depth Estimation Via Knowledge Distillation and Uncertainty Guidance 1541
Huawei Sun, Nastassia Vysotskaya, Tobias Sukianto, Hao Feng, Julius Ott, Xiangyuan Peng, Lorenzo Servadei, Robert Wille

Covariance Change Point Detection for Graph Signals 1546
Even Matencio, Charles Truong, Laurent Oudre

Robust Kernel Sparse Subspace Clustering 1551
Ivica Kopriva

Learning Binary-Antithetical Information Bottleneck for Generalizable Face Anti-Spoofing 1556
Hao Yu, Haoyu Chen, Guoying Zhao

Learned Video Compression with Refined Adaptive Flow Pyramid and Coordinate-Aware Attention 1561
Qian Huang, Wenting Liu, Xin Li, Yiming Wang

SCNN: Spike Coupling Neural Network for Multimodal Brain Network Analysis 1566
Shaolong Wei, Jiashuang Huang, Mingliang Wang, Shu Jiang, Weiping Ding

Efficient Dataset Distillation Through Low-Rank Space Sampling 1571
Hangyang Kong, Wenbo Zhou, Xuxiang He, Xiaotong Tu, Xinghao Ding

Spatially-Variant Blur Degradation Model Based on Depth Estimation 1576
Rui Xie, Shuzhan Guo, Li Zou, Jiaxiong Liu, Qian Wang, Jun Zhou

Superpoints Guided Local Explanation for Deep 3D Trackers 1581
Riran Cheng, Xupeng Wang, Ferdous Sohel, Hang Lei

HFLR: Optimizing GNN Training Via High-Fixed-Low-Resampling 1586
Chang Gong, Boyu Yang, Weiguo Zheng

GEMD-UNet: Graph Structure Enhanced Multi-Dimensional Learning Unet for Cloud Detection 1591
Jianing Chen, Chuhao Chen, Junze Yang, Wei Li, Rahul Yadav, Wenqi Zheng

Robust CLIP-Guided Deep Thinking: A Two-Stage Optimization Strategy for Enhancing Adversarial Robustness and Reliability in LVLMs 1596
Yize Sui, Wanrong Huang, Wenjing Yang, Chaofan Zhao, Jing Ren, Ji Wang

DETECLAP: Enhancing Audio-Visual Representation Learning with Object Information 1601
Shota Nakada, Taichi Nishimura, Hokuto Munakata, Masayoshi Kondo, Tatsuya Komatsu

Prompt Crossing: Evaluating Whether LLM Response Stem from Jailbreak Or Normal Prompt 1606
Kyungho Kim, Jaejin Seo, Seongmin Park, Jihwa Lee

Stance Detection for Social Text: Inference-Enhanced Multi-Task Learning with Machine-Annotated Supervision 1611
Qiankun Pi, Jicang Lu, Yepeng Sun, Qinlong Fan, Xukun Zhou, Shouxin Shang

Controlling the Number of Sample-Contributive Vertices in Generalized Sampling of Graph Signals 1616
Keitaro Yamashita, Kazuki Naganuma, Shunsuke Ono

CEMSSL: Conditional Embodied Self-Supervised Learning is All You Need for High-Precision Multi-Solution Inverse Kinematics of Robot Arms 1621
Qu Weiming, Liu Tianlin, Du Jiawei, Luo Dingsheng

Sequential Posterior Sampling with Diffusion Models 1626
Tristan S.W. Stevens, Oisín Nolan, Jean-Luc Robert, Ruud J.G. Van Sloun

Adaptive Spatiotemporal Augmentation for Improving Dynamic Graph Learning 1631
Xu Chu, Hanlin Xue, Bingce Wang, Xiaoyang Liu, Weiping Li, Tong Mo, Tuoyu Feng, Zhijie Tan

CogniDual Framework: Self-Training Large Language Models Within a Dual-System Theoretical Framework for Improving Cognitive Tasks 1636
Yongxin Deng, Xihe Qiu, Xiaoyu Tan, Chao Qu, Jing Pan, Yuan Cheng, Yinghui Xu, Wei Chu

Disentangled Representation Learning for Chinese Handwriting Recognition 1641
Tianqi Zhao, Liangrui Peng, Gang Yao, Di Wu, Yao Tao

Whisper in Medusa's Ear: Multi-Head Efficient Decoding for Transformer-Based ASR 1646
Yael Segal-Feldman, Aviv Shamsian, Aviv Navon, Gill Hetz, Joseph Keshet

RanDoctor: System-Level Ransomware Detection with ProbSparse Self-Attention 1651
Zhilu Wang, Peinan Li, Lingbo Zhao, Fengkai Yuan, Rui Hou, Dan Meng

Span Attention for Entity-Consistent Task-Oriented Dialogue Response Generation 1656
Jiale Chen, Xuelian Dong, Wenxiu Xie, Tao Gong, Fu Lee Wang, Tianyong Hao

On the Design of Low-Rank Differential Beamformers with Nonuniform Linear Microphone Arrays 1661
Hanchen Pei, Kang Chen, Gongping Huang, Jilu Jin, Jacob Benesty, Jingdong Chen

HiLiteMamba: A Lightweight and High-Frequency Aware Network for Single Image Super-Resolution 1666
Zijing Zhang, Jianfei Xiao, Bate Liu

Optimizing Speech Multi-View Feature Fusion Through Conditional Computation 1671
Weiqiao Shan, Yuhao Zhang, Yuchen Han, Bei Li, Xiaofeng Zhao, Yuang Li, Min Zhang, Hao Yang, Tong Xiao, Jingbo Zhu

Design of Robust Differential Beamformers with Microphone Arrays of Arbitrary Planar Geometry 1676
Kunlong Zhao, Xueqin Luo, Jilu Jin, Gongping Huang, Jingdong Chen, Jacob Benesty

EagerLog: Active Learning Enhanced Retrieval Augmented Generation for Log-Based Anomaly Detection 1681
Chiming Duan, Tong Jia, Yong Yang, Guiyang Liu, Jinbu Liu, Huxing Zhang, Qi Zhou, Ying Li, Gang Huang

DDNet: Exploring Dual Dependencies for Long-Term Time Series Forecasting 1686
Zihang Guo, Zijian Li, Zhiyong Yang, Zhenping Mou, Jieru Guo

Evaluating Contrastive Methodologies for Music Representation Learning Using Playlist Data 1691
Gregor Meehan, Johan Pauwels

Textual and Visual Prompt Fusion for Image Editing Via Step-Wise Alignment 1696
Zhanbo Feng, Zenan Ling, Xinyu Lu, Ci Gong, Feng Zhou, Wugedele Bao, Jie Li, Fan Yang, Robert C. Qiu

Curriculum Contrastive Learning for Aspect-Based Sentiment Analysis 1701
Zhongquan Jian, Daihang Wu, Xiangjian Zeng, Junfeng Yao, Meihong Wang, Qingqiang Wu

Rapid Online Bayesian Learning for Deep Receivers 1706
Yakov Gusakov, Osvaldo Simeone, Tirza Routtenberg, Nir Shlezinger

Self-Incremental Training for Personalized Voice Command Recognition in a Wireless Audio Sensor Network 1711
Manuele Rusci, Hugo Van Hamme, Tinne Tuytelaars

AudioComposer: Towards Fine-Grained Audio Generation with Natural Language Descriptions 1716
Yuanyuan Wang, Hangting Chen, Dongchao Yang, Zhiyong Wu, Xixin Wu

Distill to Detect: Amplifying Anomalies in Backdoor Models Through Knowledge Distillation 1721
Chang Hu, Xuyang Teng, Wenpeng Xing, Han Chen, Chenhao Ye, Meng Han

ICIMG-Net: Inject Context Information to Motion Generation for Optical Flow Estimation 1726
Wenbo Yin, Congxuan Zhang, Zhen Chen, Cheng Feng, Liyue Ge, Zige Wang

Unbiased Multimodal Audio-To-Intent Recognition 1731
Qian Dong, Yuezhou Dong, Ke Qin, Guiduo Duan, Tao He

LNLFace: Enhanced Blind Face Restoration with Local and Non-Local Lookups 1736
Weidan Yan, Wenze Shao, Dengyin Zhang

Multi-Modal Salient Object Detection Via a Unified Diffusion Model 1741
Shuo Zhang, Jiaming Huang, Wenbing Tang, Lili Tian, Yuang Wei, Jing Liu

When Sparse Graph Representation Learning Falls into Domain Shift: Feature Augmentation for Cross-Domain Graph Meta-Learning 1746
Simin Niu, Xun Liang, Sensen Zhang, Zhiyu Li, Xuan Zhang, Wu Bo, Hanyu Wang, Shichao Song, Mengwei Wang, Jiawei Yang

Radiation and Directivity Analysis of a Vibrating Dome-Shaped Radiator Mounted on an Infinite Baffle 1751
Junqing Zhang, Wen Zhang, Jingdong Chen, Jacob Benesty

Palm-Vein Images Reconstruction Against Adversarial Attacks 1756
Lunke Fei, Jiacheng Yang, Wai-Keung Wong, Shuping Zhao, Anne Toomey, Jiehang Deng

DCIM-AVSR: Efficient Audio-Visual Speech Recognition Via Dual Conformer Interaction Module 1761
Xinyu Wang, Haotian Jiang, Haolin Huang, Yu Fang, Mengjie Xu, Qian Wang

Disentangle Heart Rate Signals for Improved Stress Detection 1766
Pin-Jhao Chen, Woan-Shiuan Chien, Chi-Chun Lee

A Hybrid Probabilistic-Deterministic Model Recursively Enhancing Speech 1771
Tomohiro Nakatani, Naoyuki Kamo, Marc Delcroix, Shoko Araki

SocialRecNet: A Multimodal LLM-Based Framework for Assessing Social Reciprocity in Autism Spectrum Disorder 1776
Xin-Yu Chen, Yu-Ming Chen, Chin-Po Chen, Bo-Hao Su, Susan Shur-Fen Gau, Chi-Chun Lee

Mouth Articulation-Based Anchoring for Improved Cross-Corpus Speech Emotion Recognition 1781
Shreya G. Upadhyay, Ali N. Salman, Carlos Busso, Chi-Chun Lee

A Dynamic Edge-Selection Mechanism in HRV Hypergraph Learning for Improved Stress Detection 1786
Jing-Chun Wang, Woan-Shiuan Chien, Chi-Chun Lee

CycleFlow: Leveraging Cycle Consistency in Flow Matching for Speaker Style Adaptation 1791
Ziqi Liang, Xulong Zhang, Chang Liu, Xiaoyang Qu, Weifeng Zhao, Jianzong Wang

Toward Zero-Shot Speech Emotion Recognition Using LLMs in the Absence of Target Data..... 1796
Su Bo-Hao, Shreya G. Upadhyay, Lee Chi-Chun

Component-Wise Self-Correction Network for Human Motion Prediction 1801
Jinkai Li, Jinghua Wang, Xin Wang, Liang Yan, Yong Xu

MathReader : Text-To-Speech for Mathematical Documents 1806
Sieun Hyeon, Kyudan Jung, Nam-Joon Kim, Hyun Gon Ryu, Jaeyoung Do

A Progressive Local Variance-Guided Strategy for Improving Data Augmentation Reliability.....1811
Zheyuan Wang, Ziyao Meng, Dezhi Wu, Haoran Liao, Tianyi Wang, Hao Shen, Jiajia Li, Haitao Song

Mask Augmentation for Tumor Classification in Medical Images..... 1816
Chun-Chieh Weng, Huan-Yu Chen, Jing-Tong Tzeng, Ching-Heng Lin, Po-Chih Kuo, Chi-Chun Lee

Valve Token Masked Autoencoder for Missing Recordings on Cardiac Abnormality Classification..... 1821
An-Yan Chang, Jing-Tong Tzeng, Huan-Yu Chen, Chun-Hsiang Huang, Edward Pei-Chuan Huang, Chi-Chun Lee

Unsupervised UAV 3D Trajectories Estimation with Sparse Point Clouds..... 1826
Hanfang Liang, Yizhuo Yang, Jinming Hu, Jianfei Yang, Fen Liu, Shenghai Yuan

Two-Stream Semantic Alignment Networks for Multi-Label Image Classification..... 1831
Wenlan Kuang, Zhixin Li

FreeLesion: Synthetic Image-Mask Pairs for Fundus Lesion Segmentation Via Curriculum Learning and Feature-Loss Guided Filtering..... 1836
Peilei Fu, Song Guo

FDDSGCN: Fractional Decoupling Dynamic Spatiotemporal Graph Convolutional Network for Traffic Forecasting 1841
Jinpeng Xu, Chunna Zhao, Jing Yang, Yaqun Huang, Yaoyuan Yang, Por Lip Yee

Causal Debiasing for Visual Commonsense Reasoning..... 1846
Jiayi Zou, Gengyun Jia, Bing-Kun Bao

ADC: Enhancing Function Calling Via Adversarial Datasets and Code Line-Level Feedback..... 1851
Wei Zhang, Yi Zhang, Li Zhu, Qianghuai Jia, Feijun Jiang, Hongcheng Guo, Zhoujun Li, Mengping Zhou

Dynamic Graph Multi-Granularity Attribute Scene Evolution Sequence Recommendation 1856
Longtao Wang, Qingtian Zeng, Guiyuan Yuan, Hua Duan, Cheng Cheng, Kai Jiang

Multi-Modal Rhythmic Generative Model for Chinese Cued Speech Gestures Generation..... 1861
Li Liu, Wentao Lei, Wenwu Wang

Teaching Others Teaches Yourself: Semi-Supervised Ensembled Pseudo-Labeling Method for Image Classification 1866
Wentao Lei, Li Liu

Hypergradient-Free Training for Deep Equilibrium Models 1871
Yuhan Lin, Shengxiang Deng, Xudong Li

Heterogeneous Graph Dual-Structure Optimization Based Attribute-Aware for Recommendation 1876
Longtao Wang, Qingtian Zeng, Guiyuan Yuan, Hua Duan, Cheng Cheng, Zilong Wang

Cross-Talk Detection in the IVAS Stereo Codec Based on GCC-PHAT 1881
Vladimir Malenovsky, Tommy Vaillancourt, Milan Jelinek, Eleni Fotopoulou, Emmanuel Ravelli

Deep Receiver for Multi-Layer Data Transmission with Superimposed Pilots 1886
Jian Zou, Jian Xiao, Qingyu Mao, Shuai Liu, Bohuai Xiao, Yongsheng Liang

Retrieval-Augmented Multilingual Citation Generation 1891
Xun Liang, Simin Niu, Sensen Zhang, Zhiyu Li, Xuan Zhang, Bo Wu, Feiyu Xiong, Bo Tang, Hanyu Wang, Shichao Song, Mengwei Wang, Jiawei Yang

Subdomain Uncertainty Optimization for Cross-Speed Fault Diagnosis 1896
Jianbo Zheng, Lida Huang, Tairui Zhang, Bin Jiang, Chao Yang

Structured Random Model for Fast and Robust Phase Retrieval 1901
Zhiyuan Hu, Julián Tachella, Michael Unser, Jonathan Dong

Comparing Self-Supervised Learning Models Pre-Trained on Human Speech and Animal Vocalizations for Bioacoustics Processing 1906
Eklavya Sarkar, Mathew Magimai.-Doss

Sparse Bayesian Network for Fast Micro-Doppler Analysis 1911
Jiongge Zhang, Hang Dong, Long Tian, Xiongpeng He, Huimin Sun, Yuan Liu

Improving Micro-Expression Recognition Using Multi-Sequence Driven Face Generation 1916
Yuan Chen, Chongju Zhong, Pinyi Huang, Wangyang Cai, Lei Wang

Device-Aware Optical Adversarial Attack for a Portable Projector-Camera System 1921
Ning Jiang, Yanhong Liu, Dingheng Zeng, Yue Feng, Weihong Deng, Ying Li

Robust Identifiability for Symbolic Recovery of Differential Equations 1926
Hillary Hauger, Philipp Scholl, Gitta Kutyniok

MPOT: Manifold Preserving Optimal Transport for Visual Recognition Under Severe Distribution Shift 1931
You-Wei Luo, Zhi-Hao Li, Chuan-Xian Ren

Deep Sylvester Posterior Inference for Adaptive Compressed Sensing in Ultrasound Imaging 1936
Simon W. Penninga, Hans van Gorp, Ruud J.G. van Sloun

S-KEY: Self-Supervised Learning of Major and Minor Keys from Audio 1941
Yuexuan Kong, Gabriel Meseguer-Brocal, Vincent Lostanlen, Mathieu Lagrange, Romain Hennequin

AMSER: Accelerate Mobile Speech Emotion Recognition with Signal Compression 1946
Yu Lu, Ran Wang, Dian Ding, Han Zhang, Liyun Zhang, Lanqing Yang, Yi-Chao Chen, Guangtao Xue

ERGNN: Spectral Graph Neural Network with Explicitly-Optimized Rational Graph Filters 1951
Guoming Li, Jian Yang, Shangsong Liang

Uncovering the Visual Contribution in Audio-Visual Speech Recognition 1956
Zhaofeng Lin, Naomi Harte

FlowSep: Language-Queried Sound Separation with Rectified Flow Matching 1961
Yi Yuan, Xubo Liu, Haohe Liu, Mark D. Plumbley, Wenwu Wang

Learning Structured Compressed Sensing with Automatic Resource Allocation 1966
Han Wang, Eduardo Pérez, Iris A. M. Huijben, Hans van Gorp, Ruud Van Sloun, Florian Römer

TRACE: A Robust Framework for Malicious Traffic Detection with Noisy Labels 1971
Yitong Cai, Chengwei Peng, Shu Li, Yuyi Liu, Hongfei Zhang, Binxing Fang

Is it Still Fair? Investigating Gender Fairness in Cross-Corpus Speech Emotion Recognition 1976
Shreya G. Upadhyay, Woan-Shiuan Chien, Chi-Chun Lee

Sound-VECaps: Improving Audio Generation with Visually Enhanced Captions 1981
Yi Yuan, Dongya Jia, Xiaobin Zhuang, Yuanzhe Chen, Zhuo Chen, Yuping Wang, Yuxuan Wang, Xubo Liu, Xiyuan Kang, Mark D. Plumbley, Wenwu Wang

Lightweight Self-Supervised Monocular Depth Estimation for All-Day Scenes Using Generative Adversarial Network 1986
Junding Zhang, Di Rao, Youssef Akoudad, Wei Gao, Jie Chen

Exploration of Sequence-Wise Optimized Parameters for Low Complexity Enhancement Video Coding (LCEVC) on 4K Content 1991
Martin Benjak, Jörn Ostermann

A Federated Learning Network Intrusion Detection System for Multiple Imbalances 1996
Yifan Zeng, Peijia Zheng, Jian Li

Learning Source Disentanglement in Neural Audio Codec 2001
Xiaoyu Bie, Xubo Liu, Gaël Richard

Efficient Long Speech Sequence Modelling for Time-Domain Depression Level Estimation 2006
Shuanglin Li, Zhijie Xie, Syed Mohsen Naqvi

Enhancing Multimodal Sentiment Analysis for Missing Modality Through Self-Distillation and Unified Modality Cross-Attention 2011
Yuzhe Weng, Haotian Wang, Tian Gao, Kewei Li, Shutong Niu, Jun Du

CFSum: A Transformer-Based Multi-Modal Video Summarization Framework with Coarse-Fine Fusion 2016
Yaowei Guo, Jiazheng Xing, Xiaojun Hou, Shuo Xin, Juntao Jiang, Demetri Terzopoulos, Chenfanfu Jiang, Yong Liu

Hybrid Precoding in mmWave Multiuser MIMO Systems with Delay Alignment Modulation (DAM) 2021
Priyanka Maity, Monali Chakraborty, Suraj Srivastava, Aditya K. Jagannatham

TeXBLEU: Automatic Metric for Evaluate LaTeX Format 2026
Kyudan Jung, Nam-Joon Kim, Hyun Gon Ryu, Sieun Hyeon, Seung-Jun Lee, Hyuk-Jae Lee

AGR: Age Group Fairness Reward for Bias Mitigation in LLMs 2031
Shuirong Cao, Ruoxi Cheng, Zhiqiang Wang

A Critical Assessment of Visual Sound Source Localization Models Including Negative Audio 2036
Xavier Juanola, Gloria Haro, Magdalena Fuentes

A Unified Model for Oral Reading Fluency and Student Prosody 2041
Yihao Wang, Zhongdi Wu, Joseph Nese, Akihito Kamata, Vedant Nilabh, Eric C. Larson

Continuous-Discrete Differentiable Particle Filters for Irregular Time Series 2046
Hao Wen, Paul Krause, Lee Gillam

Speech Few-Shot Learning for Language Learners' Speech Recognition 2051
Jian Cheng, Sam Nguyen

Denoising Student Features with Diffusion Models for Knowledge Distillation in Speaker Verification 2056
Zezhong Jin, Youzhi Tu, Zhe Li, Zilong Huang, Chong-Xin Gan, Man-Wai Mak

Improved Out-Of-Domain Detection in VAE Latent Spaces with Boundary-Driven Regularisation 2061
Miao Jing, Vidhyasaharan Sethu, Beena Ahmed

Evidential Neural GPLDA: A Novel Approach to Quantify Prediction Uncertainty in Speaker Verification Systems 2066
Miao Jing, Vidhyasaharan Sethu, Beena Ahmed

Advancing High-Resolution and Efficient Automotive Radar Imaging Through Domain-Informed 1D Deep Learning 2071
Ruxin Zheng, Shunqiao Sun, Hongshan Liu, Holger Caesar, Honglei Chen, Jian Li

SLiCK: Exploiting Subsequences for Length-Constrained Keyword Spotting 2076
Kumari Nishu, Minsik Cho, Devang Naik

Unrolled Generative Compound Gaussian Network for Computer Tomography 2081
Carter Lyons, Raghu G. Raj, Margaret Cheney

LFSRDiff: Light Field Image Super-Resolution Via Diffusion Models 2086
Wentao Chao, Junli Zhao, Fuqing Duan, Guanghui Wang

MVCBRec: Multi-View Contrastive Learning for Bundle Recommendation 2091
Mengmeng Li, Jinlong Tian, Yongqiang Zhao, Hongmei Li, Xudong Fang

FedRPN: An Efficient Framework for Optimizing System Heterogeneity in Federated Learning 2096
Baolu Xue, Hanyuan Zheng, Jiale Zhang, Jiewen Liu, Bing Chen

LAVViT: Latent Audio-Visual Vision Transformers for Speaker Verification 2101
R. Gnana Praveen, Jahangir Alam

Training Dialogue Systems by AI Feedback for Improving Overall Dialogue Impression 2106
Kai Yoshida, Masahiro Mizukami, Seiya Kawano, Canasai Kruengkrai, Hiroaki Sugiyama, Koichiro Yoshino

Refiner: Fine-Grained Cross-Modal Concepts Refinement for Compositional Zero-Shot Learning 2111
Xiao Zhang, Haodong Jing, Hui Chen, Yongqiang Ma, Nanning Zheng

Adaptive Aspect Ratios with Patch-Mixup-ViT-Based Vehicle ReID 2116
Mei Qiu, Lauren Ann Christopher, Stanley Chien, Lingxi Li

ASFC-NeRF: Large-Scale Scene Rendering with Adaptive Sampling and Feature-Aware Compression 2121
Xinrui Zhang, Yufeng Wang, Shuangkang Fang, Zesheng Wang, Huayu Zhang, Dacheng Qi, Wenrui Ding

Advancing Streaming ASR with Chunk-Wise Attention and Trans-Chunk Selective State Spaces 2126
Masato Mimura, Takafumi Moriya, Kohei Matsuura

30+ Years of Source Separation Research: Achievements and Future Challenges 2131
Shoko Araki, Nobutaka Ito, Reinhold Haeb-Umbach, Gordon Wichern, Zhong-Qiu Wang, Yuki Mitsufuji

An Efficient Residual-Based Low-Dose PET Reconstruction with Spatial-Frequency Integration 2136
Minghui Li, Lei Yu, Hewen Pan, Shengqing Hu, Longling Zhang, Shengshan Hu, Wei Wan, Peijin Guo

Massive MIMO System Partitioning for Efficient Hybrid Beamformer Optimization 2141
Saidur R. Pavel, Yimin D. Zhang, Batu K. Chalise

MPFL: A Decentralised Federated Learning Framework Based on Multi-Population Genetic Algorithm 2146
Wenqi Ding, Yuanchao Liu, Zhongjie Wang, Zheng Chu

2.5D Top-K Ranked Multiple Instance Learning to Classify NSCLC PD-L1 Status on CT Images 2151
Rui Xu, Huadong Liu, Yongcen Li, Xinchen Ye, Zhihui Wang, Hongkai Wang, Haojie Li, Yi Wang, Dingpin Huang, Fangyi Xu, Yi Gan, Yuan Tu, Hongjie Hu

Enhanced Multimodal Depression Detection with Emotion Prompts 2156
Shiyu Teng, Jiaqing Liu, Hao Sun, Shurong Chai, Tomoko Tateyama, Lanfen Lin, Yen-Wei Chen

MTDA-HSED: Mutual-Assistance Tuning and Dual-Branch Aggregating for Heterogeneous Sound Event Detection 2161
Zehao Wang, Haobo Yue, Zhicheng Zhang, Da Mu, Jin Tang, Jianqin Yin

ADD: A Detection Method for Image-Processing Adversarial Defenses 2166
Da Zhang, Jiazheng Sun, Chenxiao Xia, Ruinan Ma, Jun Zheng

Shifting Spotlight for Co-Supervision: A Simple Yet Efficient Single-Branch Network to See Through Camouflage 2171
Yang Hu, Jinxia Zhang, Kaihua Zhang, Yin Yuan, Jiale Huang, Zechao Zhan, Xin Wang

Fioma: Towards Open-Set Semi-Supervised Specific Emitter Identification 2176
Qingyun Xu, Lixiang Liu, Xin Zhou

MSACC: A Unified Multimodal Sentiment Analysis Framework for High Interpretability and Zero-Shot Performance 2181
Yi Liang, Turdi Tohti, Bo Kong, Dongfang Han, Tianwei Yan, Askar Hamdulla

Diagram Formalization Enhanced Multi-Modal Geometry Problem Solver 2186
Zeren Zhang, Jo-Ku Cheng, Jingyang Deng, Lu Tian, Jinwen Ma, Ziran Qin, Xiaokai Zhang, Na Zhu, Tuo Leng

Adaptive Lossless Compression for Genomics Data by Multiple (s, K)-Mer Encoding and XLSTM 2191
Hui Sun, Yanfeng Ding, Liping Yi, Huidong Ma, Haonan Xie, Gang Wang, Xiaoguang Liu

Automatic Numbering and Pathological Recognition of Pediatric Teeth Using CNN and Attention Mechanisms 2196
Hongzhou Zhu, Yuhao Qiu, Renjie Hu, Ang Li, Shengji Zhu, Lei Wang

A Study on Zero-Shot Non-Intrusive Speech Assessment Using Large Language Models 2201
Ryandhimas E. Zezario, Sabato M. Siniscalchi, Hsin-Min Wang, Yu Tsao

Semi-Supervised Knowledge Distillation Framework Towards Lightweight Large Language Model for Spoken Language Translation 2206
Tonmoy Rajkhowa, Amartya Roy Chowdhury, Achyut Mani Tripathi, Sanjeev Sharma, Om Jee Pandey

FashionFAE: Fine-Grained Attributes Enhanced Fashion Vision-Language Pre-Training 2211
Jiale Huang, Dehong Gao, Jinxia Zhang, Zechao Zhan, Yang Hu, Xin Wang

Soft Knowledge Distillation with Multi-Dimensional Cross-Net Attention for Image Restoration Models Compression 2216
Yongheng Zhang, Danfeng Yan

Fast Inter-Frame Coding for Dynamic Meshes Via Supervoxel-Based Shape Matching 2221
Xudong Jin, Jianfeng Xu, Kei Kawamura

Integrating Potential Pronunciations for Enhanced Mispronunciation Detection and Diagnosis Ability in LLMs 2226
Minglin Wu, Jing Xu, Xueyuan Chen, Helen Meng

OPFormer: Real-Time Optimal Power Flow with CNN-Based Transformer 2231
Kaijie Xu, Xilin Dai, Lin Qiu

Stealthy Backdoor Attack Against Video Recognition Models 2236
Jiale Yan, Bo Zhao, Chunyu Yang

Salmon: A Suite for Acoustic Language Model Evaluation 2241
Gallil Maimon, Amit Roth, Yossi Adi

Large Language Model Should Understand Pinyin for Chinese ASR Error Correction 2246
Yuang Li, Xiaosong Qiao, Xiaofeng Zhao, Huan Zhao, Wei Tang, Min Zhang, Hao Yang

Quantum Multi-Path Communication Protocol Based on Maximum Flow Theory 2251
Jihao Fan, Liang Huang, Jun Li, Long Shi, Yuwen Qian

DH-VTON: Deep Text-Driven Virtual Try-On Via Hybrid Attention Learning 2256
Jiabao Wei, Zhiyuan Ma

Automatic Detection of Domain Shifts in Speech Enhancement Systems Using Confidence-Based Metrics 2261
Lior Frankel, Shlomo E. Chazan, Jacob Goldberger

MambaRF: A Bi-Directional Mamba Structure for Radio Frequency Signal Classification of Unmanned Aerial Vehicle 2266
Mufeng Yao, Chao Liu, Lexu Xie, Mingmin Chi

Movable Antenna Aided Physical Layer Security with No Eavesdropper CSI 2271
Zhenqiao Cheng, Chongjun Ouyang, Xingqi Zhang

Foundation Models Boost Low-Level Perceptual Similarity Metrics 2276
Abhijay Ghildyal, Nabajeet Barman, Saman Zadtootaghaj

Gaussian-Face: Talking Head Generation with Hybrid Density Via 3D Gaussian Splatting 2281
Guanwen Feng, Yilin Zhang, Yunan Li, Siyu Jin, Qiguang Miao

Advancing Paired Image-Mask Synthesis for Automated Nanoparticle Phenotyping 2286
Xiaoqin Tang, Chaohui Liu, Guoqiang Xiao

Seg-Diffusion: Text-To-Image Diffusion Model for Open-Vocabulary Semantic Segmentation 2291
Shuo Zhang, Jiaming Huang, Yan Wu, Tao Hu, Wenbing Tang, Jing Liu

A Novel Decision-Making Model for Playing Board Game Combining Planning and Opponent Behaviors 2296
Jiajing Zhang, Jiamei Jiang, Linjing Li, Daniel Zeng

A Cross-Modal Multi-Attitude Framework for the Generation of Space Target ISAR Images 2301
Derong Kong, Huaizhang Liao, Jingyuan Xia

Speaker-IPL: Unsupervised Learning of Speaker Characteristics with i-Vector Based Pseudo-Labels 2306
Zakaria Aldeneh, Takuya Higuchi, Jee-Weon Jung, Li-Wei Chen, Stephen Shum, Ahmed Hussen Abdelaziz, Shinji Watanabe, Tatiana Likhomanenko, Barry-John Theobald

Sign-Mamba: Advanced Mamba-Based Sign Language Generation 2311
Guanwen Feng, Yilin Zhang, Yunan Li, An Liu, Qiguang Miao

Anima2: Cross-Species Animal Animation Through Image-To-Video Synthesis with Subject Alignment 2316
Yuanfeng Xu, Yuhao Chen, Zhongzhan Huang, Zijian He, Guangrun Wang, Liang Lin

Learning-Based Utility Estimation with Application to Speech Enhancement of a Moving Speaker 2321
Jie Zhang, Chengqian Jiang, Yichi Wang, Haoyin Yan, Miao Sun

AdaptVC: High Quality Voice Conversion with Adaptive Learning 2326
Jaehun Kim, Ji-Hoon Kim, Yeunju Choi, Tan Dat Nguyen, Seongkyu Mun, Joon Son Chung

Medical Image Segmentation Via Sparse Coding Decoder 2331
Long Zeng, Mingwei Zhu, Kaigui Wu, Zefang Li

KAN-Face: Efficient Resource Usage and Precision Lip-Sync in Talking Head Generation 2336
Guanwen Feng, Siyu Jin, Zhihao Qian, Yunan Li, Qiguang Miao

Towards Automatic Assessment of Self-Supervised Speech Models Using Rank 2341
Zakaria Aldeneh, Vimal Thilak, Takuya Higuchi, Barry-John Theobald, Tatiana Likhomanenko

PEDE: Enhance Multi-Modal Sarcasm Detection in Videos Via Prompted Emotion Distributions 2346
Xiaoqiang Zhang, Ying Chen, Guangyao Li, Buwen Liang

Self-Supervised Monocular Depth Estimation from Videos Via Pose-Adaptive Reconstruction 2351
Xin Sun, Boqian Liu, Xinchen Ye, Rui Xu, Haojie Li

R2-SAC: A Relaxation-And-Refinement SAC Agent for Stock Portfolio Trading 2356
Xiaoyun Han, Jun Wang

Transfer Learning Via Functional Balancing in Reproducing Kernel Hilbert Spaces 2361
Boyan Gu, Sheng Zheng, Xiaojun Mao, Zhonglei Wang

Blind Estimation of Sub-Band Acoustic Parameters from Ambisonics Recordings Using Spectro-Spatial Covariance Features 2366
Hanyu Meng, Jeroen Breebaar, Jeremy Stoddard, Vidhyasaharan Sethu, Eliathamby Ambikairajah

Few-Shot Image Classification Based on Attribute Prediction and Selection 2371
Xin Sun, Boqian Liu, Xinchen Ye, Guanqiao Chen, Rui Xu, Haojie Li

Edge-Aware Laplacian Pyramid Network for Efficient Image Deblurring 2376
Zhe Xu, Zhipei Lei, Dingyong Gou, Yanlin Wu, Liwen Zhang, Cong Li

Automatic Labelling & Semantic Segmentation with 4D Radar Tensors..... 2381
Botao Sun, Ignacio Roldan, Francesco Fioranelli

Less Yet Robust: Crucial Region Selection for Scene Recognition..... 2386
Jianqi Zhang, Mengxuan Wang, Jingyao Wang, Lingyu Si, Changwen Zheng, Fanjiang Xu

Networked ISAC Beamforming Design with Capacity-Limited Fronthaul Links 2391
Kexin Zhang, Yanqing Xu, Tsung-Hui Chang

Mining Scene Structural Guidance for Thermal Images in Self-Supervised Monocular Depth Estimation..... 2396
Xinchen Ye, Xia Mao, Rui Xu, Haojie Li

A Dual-Perspective Metaphor Detection Framework Using Large Language Models..... 2401
Yujie Lin, Jingyao Liu, Yan Gao, Ante Wang, Jinsong Su

Multi-Point Positional Insertion Tuning for Small Object Detection 2406
Kanoko Goto, Takumi Karasawa, Takumi Hirose, Rei Kawakami, Nakamasa Inoue

Delving into Transformer-Based Network Architecture for Guided Depth Super-Resolution 2411
Xinchen Ye, Aokai Zhang, Rui Xu, Haojie Li

Do Multimodal Language Models Really Understand Direction? a Benchmark for Compass Direction Reasoning 2416
Hang Yin, Zhifeng Lin, Xin Liu, Bin Sun, Kan Li

Complex Coprime Frequency Sum Based Signal Representation for Period Estimation..... 2421
Shaik Basheeruddin Shah, Nazar T. Ali, Ahmed Altunaiji, Vijay Kumar Chakka, Mohamed Alhajri

UV-Mamba: A DCN-Enhanced State Space Model for Urban Village Boundary Identification in High-Resolution Remote Sensing Images 2426
Lulin Li, Ben Chen, Xuechao Zou, Junliang Xing, Pin Tao

Post-Hoc Adversarial Stickers Against Micro-Expression Leakage..... 2431
Pei-Sze Tan, Sailaja Rajanala, Yee-Fan Tan, Arghya Pal, Chun-Ling Tan, Raphaël C.-W. Phan, Huey-Fang Ong

Robust Cross-Etiology and Speaker-Independent Dysarthric Speech Recognition..... 2436
Satwinder Singh, Qianli Wang, Zihan Zhong, Clarion Mendes, Mark Hasegawa-Johnson, Waleed Abdulla, Seyed Reza Shahamiri

Virtual Leader-Based Safe Formation-Switching Control for Dense Environments..... 2441
Aamna Zahid Piracha, Bernhard Rinner

Enhancing Multi-Channel Speech with Limited Microphones Via Spherical Harmonic Transform 2446
Jiahui Pan, Hui Zhang, Xueliang Zhang

M2PAIR: A High-Quality Acoustic Impulse Response Computation Model 2451
Zhiyu Li, Xinpei Zhao, Jing Wang, Xinyuan Qian, Xiang Xie

AVS3P10 Standard for Real-Time Speech Coding 2456
Wei Xiao, Weibei Dou, Wenlong Wang, Gaoxiong Yi, Jingxin Li, Shidong Shang

Improving GAN Performance Using Confidence-Aware Discrimination 2461
Jinfeng Wu, Wu Shi

Hop-Level Direct Preference Optimization for Knowledge Graph Reasoning with Trees 2466
Tiesunlong Shen, Jin Wang, Xuejie Zhang, Erik Cambria

Windowed Quantum Phase Estimation: Signal Processing Approach to a Quantum Algorithm 2471
Xue Wen

BCG Data Imputation Via Multimodal Feature Alignment and Semantic Sequence Prediction 2476
Jiafeng Qiu, Huadan Wang, Peihan Yao, Gang Shen

Steering Large Language Models for Vulnerability Detection 2481
Jiayuan Li, Lei Cui, Jie Zhang, Haiqiang Fei, Yu Chen, Hongsong Zhu

AdaCS: Adaptive Normalization for Enhanced Code-Switching ASR 2486
The Chuong Chu, Vu Tuan Dat Pham, Trung Kien Dao, Ngoc Hoang Nguyen, Steven Truong

Inside and Inside: Efficient Anomaly Detection by Fully Capturing the Detailed Dynamics 2491
Ziyu Tang, Xiren Zhou, Ao Chen, Shikang Liu, Chuyang Wei, Huanhuan Chen

Towards Dynamic Skeleton-Based Handshape Subunits for Sign Language Assessment 2496
Sandrine Tornay, Mathew Magimai-Doss

Enhancing Text Annotation Through Rationale-Driven Collaborative Few-Shot Prompting 2501
Jianfei Wu, Xubin Wang, Weijia Jia

COSMIC Waveforms for Integrated Communication and Imaging 2506
Marco Manzoni, Francesco Linsalata, Maurizio Magarini, Stefano Tebaldini

3D View Optimization for Improving Image Aesthetics 2511
Taichi Uchida, Yoshihiro Kanamori, Yuki Endo

Kronecker-Structured Sparse Vector Recovery with Application to IRS-MIMO Channel Estimation 2516
Yanbin He, Geethu Joseph

Single-Loop Variance-Reduced Stochastic Algorithm for Nonconvex-Concave Minimax Optimization 2521
Xia Jiang, Linglingzhi Zhu, Taoli Zheng, Anthony Man-Cho So

RefCap: Zero-Shot Video Corpus Moment Retrieval Based on Refined Dense Video Captioning 2526
Yi Pan, Yujia Zhang, Michael Kampffmeyer, Xiaoguang Zhao

DOA Estimation Based on Enhanced SRP-MVDR Using Kronecker Product Decomposition for Large Rectangular Microphone Arrays 2531
Yichen Zeng, Jilu Jin, Gongping Huang, Jingdong Chen, Jacob Benesty

Vision Mamba-Based Approach for Incomplete Boundary Document Image Rectification 2536
Weihao Zhang, Xin Xia, Maopeng Li, Yunbo Zhao

FAWL: Weakly-Supervised Video Corpus Moment Retrieval with Frame-Wise Auxiliary Alignment and Weighted Contrastive Learning 2541
Yi Pan, Yujia Zhang, Xiaoguang Zhao

Multi-Hop Self-Augmented Graph Contrastive Learning for Node Classification 2546
Yutong Wang, Xiaofeng Meng, Minhao Zou, Siyang Leng

MMCD: Memory-Based Multimodal Change Detection 2551
Limeng Zhang, Zenghui Zhang, Juanping Wu, Weiwei Guo, Tao Zhang, Wenxian Yu

Adapting Without Seeing: Text-Aided Domain Adaptation for Adapting CLIP-Like Models to Novel Domains 2556
Louis Hémadou, Héléna Vorobieva, Ewa Kijak, Frédéric Jurie

Amplitude-Guidance Low-Light Image Enhancement with Frequency-Based Channel Attention 2561
Jinlong Wang, Xiongxin Tang, Fanjiang Xu, Hanxiang Yang

CAPAST: Content Affinity Preserved Arbitrary Style Transfer 2566
Xinyuan Zheng, Xiaojie Li, Canghong Shi, Jia He, Zhan Ao Huang, Xian Zhang, Imran Mumtaz

Enhancing Session-Based Recommendation with Hypergraph Motifs and Contrastive Learning 2571
Tingxuan Chen, Liu Yang, Zidong Wang, Guohui Li, Jun Long

Smooth-Foley: Creating Continuous Sound for Video-To-Audio Generation Under Semantic Guidance 2576
Yaoyun Zhang, Xuenan Xu, Mengyue Wu

RPPFL: Robust and Privacy-Preserving Federated Learning Via Trusted Execution Environments 2581
Xiaolei Zhang, Zhaoyu Chen, Guangpu Chen, Xinyu Feng, Qingni Shen, Zhonghai Wu

Spatial-Frequency Information Interaction Diffusion for SAR Colorization 2586
Xupei Zhang, Hanlin Qin, Jingjing Li, Jinni Geng, Zihan Gao, Yue Yu

Improving Generated and Retrieved Knowledge Combination Through Zero-Shot Generation 2591
Xinkai Du, Quanjie Han, Chao Lv, Yan Liu, Yalin Sun, Hao Shu, Hongbo Shan, Maosong Sun

Decentralized Federated Dataset Dictionary Learning for Multi-Source Domain Adaptation 2596
Rebecca Clain, Eduardo Fernandes Montesuma, Fred Ngolè Mboula

RF-GML: Reference-Free Generative Machine Listener 2601
Arijit Biswas, Guanxin Jiang

SBA: A Swift and Stealthy Backdoor Attack Framework for Federated Learning 2606
Junhan Wang, Zhangming Wu, Zhuoyue Wang, Lu Dong

Why Disentanglement-Based Speaker Anonymization Systems Fail at Preserving Emotions? 2611
Ünal Ege Gaznepoglu, Nils Peters

Spatial-Temporal Perception with Causal Inference for Naturalistic Driving Action Recognition 2616
Qing Chang, Wei Dai, Zhihao Shuai, Limin Yu, Yutao Yue

Automatic Recognition of Rodent Call Types Using Deep Supervectors 2621
Fasih Haider, Raven Hickson, Peter Kind, Saturnino Luz

Explaining Representations in Correlation-Based Deep Multiview Representation Learning 2626
Maurice Kuschel, Amr Alkhatib, Tanuj Hasija, Henrik Boström

FA-GAN: Defense Against Adversarial Attacks in Automatic Modulation Recognition 2631
Shilong Zhang, Yu Song, Shubin Wang

UIDAPLE: Unsupervised Incremental Domain Adaptation Through Adaptive Prompt Learning 2636
Samrat Mukherjee, Tanuj Sur, Saurish Seksaria, Subhasis Chaudhuri, Gemma Roig, Biplab Banerjee

Enhancing Network Calibration for Low-Cost Gas Sensor Networks Through Adaptive Similarity Search 2641
Cheng Yang, Saikat Chatterjee, Tobias J. Oechtering

USD: Unsupervised Soft Contrastive Learning for Fault Detection in Multivariate Time Series 2646
Hong Liu, Xiuxiu Qiu, Yiming Shi, Zelin Zang

RSM: Refined Saliency Map for Explainable 3D Object Tracking 2651
Riran Cheng, Xupeng Wang, Ferdous Sohel, Hang Lei

Unifying Within and Across: Intra-Modality Multi-View Fusion and Inter-Modality Alignment for Knowledge Graph Completion 2656
Zhen Li, Jibin Wang, Zhuo Chen, Kun Wu, Meng Ai, Leike An, Liqiang Wang, Haoxuan Li

Hyperbolic Multimodal Knowledge Graph Embedding 2661
Qiuyu Liang, Weihua Wang, Cunda Wang, Feilong Bao, Jie Yu

GaussianEnhancer: A General Rendering Enhancer for Gaussian Splatting 2666
Chen Zou, Qingsen Ma, Jia Wang, Ming Lu, Shanghang Zhang, Zhaofeng He

Conditional Latent Diffusion-Based Speech Enhancement Via Dual Context Learning 2671
Shengkui Zhao, Zexu Pan, Kun Zhou, Yukun Ma, Chong Zhang, Bin Ma

Learning Stroke-Order Dynamics in Few-Shot Font Generation Via Sequential Awareness 2676
Jinshan Zeng, Yiyang Yuan, Yan Zhang, Yefei Wang, Xijia Wang

ChannelMixer: A Hybrid CNN-Transformer Framework for Enhanced Multivariate Long-Term Time Series Forecasting 2681
Erlei Zhang, Wenxuan Yuan, Xiangsen Liu

A Scale-Adaptive and Background-Robust Method for Surface Defect Detection 2686
Jiahao Dong, Zuo Zuo, Zongze Wu, Meiqin Liu

Attribute Conditional Diffusion-Augmented Person Re-Identification 2691
Shijie Nie, Ziqiang Shi, Rujie Liu, Song Guo, Meng Zhang, Mengjiao Wang, Kazuki Osamura, Lina Septiana, Abe Narishige

HiFi-SR: A Unified Generative Transformer-Convolutional Adversarial Network for High-Fidelity Speech Super-Resolution 2696
Shengkui Zhao, Kun Zhou, Zexu Pan, Yukun Ma, Chong Zhang, Bin Ma

Distributed Interference Alignment Precoding and Detection for MU-MIMO OTSM Downlink in Time-Varying Channels 2701
Sapta Girish Neelam

Streaming Keyword Spotting Boosted by Cross-Layer Discrimination Consistency 2706
Yu Xi, Haoyu Li, Xiaoyu Gu, Hao Li, Yidi Jiang, Kai Yu

Energy Backdoor Attack to Deep Neural Networks 2711
Hanene F. Z. Brachemi Meftah, Wassim Hamidouche, Sid Ahmed Fezza, Olivier Déforges, Kassem Kallas

No-Reference Point Cloud Quality Assessment Based on Graph Signal Variation 2716
Ryosuke Watanabe, Keisuke Nonaka, Eduardo Pavez, Tatsuya Kobayashi, Antonio Ortega

An Attribute-Enriched Dataset and Auto-Annotated Pipeline for Open Detection 2721
Pengfei Qi, Yifei Zhang, Wenqiang Li, Youwen Hu, Kunlong Bai

NeRF-VLD: Efficient Visual Landmark Database Construction Via Scene Constraints 2726
Tao Chen, Qun Niu, Ning Liu

Deep Learning-Based Perceptual Vibrotactile Codec with Rate Scalability 2731
Lars Nockenberg, Wenxuan Wei, Mariam Navai, Eckehard Steinbach

SAM Adaptation with Refocused Attention and Diverse Prompts for Medical Image Segmentation 2736
Liangshan Zhu, Xing Wu, Chengliang Wang, Haidong Wang

Semantic-Guided Gaussian Splatting with Deferred Rendering 2741
Nan Wang, Xiaohan Yan, Xiaowei Song, Zhicheng Wang

Mixed Gaussian Splatting for High-Quality Rendering and Reconstruction 2746
Shuai Liu, Jianyu Ding, Jie Yang, Wei Liu

Hybrid Feature Collaborative Reconstruction Network for Few-Shot Fine-Grained Image Classification 2751
Shulei Qiu, Wanqi Yang, Ming Yang

Diffused Poses and Distilled Expressions for Controllable Audio-Driven Talking Face Generation 2756
Ziqi Zhou, Weize Quan, Zhaojin Lu, Dong-Ming Yan

Boosting Open-Vocabulary Object Detection Performance Via Class-Agnostic Pseudo-Labels and MultiModal Hybrid Knowledge 2761
Ziyang Chen, Dongqin Liu, Jiao Dai, Songlin Hu

MLNet: Mutual Learning Network to Improve Self-Supervised Representation for Fine-Grained Visual Recognition 2766
Peipei Zhao, Jiaxuan Wang, Zixiang Lu, Qiguang Miao

Efficient Pruning for Large-Scale Seq2Seq Speech Models Without Back-Propagation 2771
Tianteng Gu, Bei Liu, Yanmin Qian

Fundamental Social Learning Scaling Law for Tracking Hidden Markov Models 2776
Malek Khammassi, Virginia Bordignon, Vincenzo Matta, Ali H. Sayed

BID-Net: Balanced Incremental Distillation Network for Fair Dermatological Disease Diagnosis 2781
Yiqin Luo, Tianlong Gu, Fengrui Hao, Liang Chang

Intrusion Detection for Intelligent Transportation Systems: A Lightweight Interpretable Model 2786
Yuxi Zhou, Tao Feng, Yazhuo Gao, Yixuan Wu, Lin Yang, Jiaqi Lin

Active Listener: Continuous Generation of Listener's Head Motion Response in Dyadic Interactions 2791
Bishal Ghosh, Emma Li, Tanaya Guha

Single-View Reconstruction Via Decoupled 3D Gaussian Splatting 2796
Sheng Liu, Shiming Zhu, Huilong Pi, Yunchuan Qin, Zhuo Tang, Ruihui Li

MpoxMamba: A Grouped Mamba-Based Lightweight Hybrid Network for Mpox Detection 2801
Yubiao Yue, Jun Xue, Haihuang Liang, Zhenzhang Li, Yufeng Wang

Automated Graph Attention Network for Heterogeneous Entity Resolution 2806
Chen Liu, Xiaohui Rong

Towards an Integrated Approach for Expressive Piano Performance Synthesis from Music Scores 2811
Jingjing Tang, Erica Cooper, Xin Wang, Junichi Yamagishi, György Fazekas

Scaling Bioacoustic Signal Pre-Training with Million Samples Via Mask-Modeling 2816
Xuyao Deng, Tianjiao Wan, Kele Xu, Tian Gao, Peng Qiao, Dawei Feng, Yong Dou

A Novel Weighted Sparse Component Analysis for Underdetermined Blind Speech Separation 2821
Yudong He, Baeck Hyun Woo, Richard H. Y. So

RQTalker: Speech-Driven 3D Facial Animation Via Region-Aware Vector Quantization 2826
Yingying Fan, Kaisiyuan Wang, Hang Zhou, Shengyi He, Yu Wu

Mamba-YOLO-World: Marrying YOLO-World with Mamba for Open-Vocabulary Detection 2831
Haoxuan Wang, Qingdong He, Jinlong Peng, Hao Yang, Mingmin Chi, Yabiao Wang

A Novel Audio-Visual Multimodal Semi-Supervised Model Based on Graph Neural Networks for Depression Detection 2836
Yaqin Li, Chenjian Sun, Yihong Dong

Density-Aware and Depth-Aware Visual Representation for Zero-Shot Object Counting 2841
Fang Nan, Feng Tian, Ni Zhang, Nian Liu, Haonan Miao, Guang Dai, Mengmeng Wang

Hybrid Feature Fusion for Enhancing Medical Document Embedding 2846
Yi Zhu, Xiangyang Liu, Tianqi Pang, Xuncan Xiao, Xiaofan Zhang, Chenyou Fan

Key Clues Guided Video Character Social Relationship Recognition Enhanced by LLM 2851
Wenlong Dong, Qing Zhu, Qirong Mao

OSR: Toward Developing Efficient Federated Learning-Based Human Activity Recognition Using Optimal Server Representations 2856
Ensieh Khazaei, Bilal Taha, Alireza Esmaeilzehi, Dimitrios Hatzinakos

MusicMamba: A Dual-Feature Modeling Approach for Generating Chinese Traditional Music with Modal Precision 2861
Jiatao Chen, Xing Tang, Tianming Xie, Jing Wang, Wenjing Dong, Bing Shi

Low-Complexity Cramér-Rao Lower Bound and Sum Rate Optimization in ISAC Systems 2866
Tianyu Fang, Nhan Thanh Nguyen, Markku Juntti

Microtitre Plate Image Augmentation with Generative Adversarial Networks 2871
Ru Li, Tingting Chai, Samaneh Kouchaki, David Clifton, Yang Yang

Towards Sub-Millisecond Latency Real-Time Speech Enhancement Models on Hearables 2876
Artem Dementyev, Chandan K. A. Reddy, Scott Wisdom, Navin Chatlani, John R. Hershey, Richard F. Lyon

On Class Separability Pitfalls in Audio-Text Contrastive Zero-Shot Learning 2881
Tiago Fernandes Tavares, Fabio José Ayres, Zhepei Wang, Paris Smaragdis

Mixed-Precision Graph Neural Quantization for Low Bit Large Language Models 2886
Wanlong Liu, Yichen Xiao, Dingyi Zeng, Hongyang Zhao, Wenyu Chen, Malu Zhang

Proud-SLAM: Neural Point-Based Hybrid RGBD Monocular Dense SLAM 2891
Wenzhi Guo, Lijun Chen

3DSignDiff: Towards 3D Sign Language Gesture Generation 2896
Ronghao Yu, Yun Liu, Xiyue Bai, Rui Yang, Yingna Wu

Exploring Triple Knowledge Cues for Zero-Shot Human-Object Interaction Detection 2901
Fang Nan, Ni Zhang, Qidong Liu, Wei Jing, Guang Dai, Yan Chen, Feng Tian

HyperDiff: Masked Diffusion Model with High-Efficient Transformer for Hyperspectral Image Cross-Scene Classification 2906
Pei Zhang, Dong Wang, Chanyue Wu, Jing Yang, Lei Kang, Zongwen Bai, Ying Li, Qiang Shen

Joint Multi-Scale Contextual and Noise Suppression for Group Emotion Recognition 2911
Wangdong Guo, Qing Zhu, Qirong Mao

Graph Topology Identification Based on Covariance Matching 2916
Yongsheng Han, Alberto Natali, Geert Leus

Multiscale Adaptive Channel Estimation for OTFS 2921
Prabhat Kumar, Chandra R. Murthy

DGJA: Dependency Graph-Enhanced Joint Attention Structure for Multimodal Sarcasm Detection 2926
Yiming Liu, Rui Song, Lida Shi, Ling Gao, Hao Xu

LKA-ReID:Vehicle Re-Identification with Large Kernel Attention 2931
Xuezhi Xiang, Zhushan Ma, Lei Zhang, Denis Ombati, Himaloy Himu, Xiantong Zhen

IOR: Inversed Objects Replay for Incremental Object Detection 2936
Zijia An, Boyu Diao, Libo Huang, Ruiqi Liu, Zhulin An, Yongjun Xu

Graph Pooling Via Dropping Task-Irrelevant Nodes 2941
Cheng Zhong, Shaofeng Zhang, Feng Zhu, Rui Zhao, Xiaokang Yang, Junchi Yan

Imperceptible Transfer Attack on Large Vision-Language Models 2946
Xiaowen Cai, Daizong Liu, Runwei Guan, Pan Zhou

Transfer Risk Map: Mitigating Pixel-Level Negative Transfer in Medical Segmentation 2951
Shutong Duan, Jingyun Yang, Yang Tan, Guoqing Zhang, Yang Li, Xiao-Ping Zhang

MusicGen-Stem: Multi-Stem Music Generation and Edition Through Autoregressive Modeling 2956
Simon Rouard, Robin San Roman, Yossi Adi, Axel Roebel

Melody Structure Transfer Network: Generating Music with Separable Self-Attention 2961
Junlin Wu, Ning Zhang, Cheng Zhong, Boan Chen, Huanxi Liu, Junchi Yan

Imperceptible Adversarial Attacks on Point Clouds Guided by Point-To-Surface Field 2966
Keke Tang, Weiyao Ke, Weilong Peng, Xiaofei Wang, Ziyong Du, Zhize Wu, Peizan Zhu, Zhihong Tian

A Novel Split Deep Unfolding Transformer for Pan-Sharpening 2971
Jiannan Chen, Zhizhuo Jiang, Xueqian Wang, Yaowen Li, Huajie Wang, Yu Liu

A Multi-Annotated and Multi-Modal Dataset for Wide-Angle Video Quality Assessment 2976
Bo Hu, Wei Wang, Chunyi Li, Lihuo He, Leida Li, Xinbo Gao

TELL ME: Tackle Electrocardiogram with Large Language Model Effectively 2981
Tianyi Shi, Siyang Zheng, Zhu Meng, Zhe Cui, Jin Huang, Changrui Ren, Bo Zhang, Zhicheng Zhao

Towards Green VAE: A Light Pixel-Weighting Technique to Enhance Variational AutoEncoder 2986
Cheng Zhong, Junlin Wu, Ziming Feng, Boan Chen, Junchi Yan

A Practical Gated Recurrent Transformer Network Incorporating Multiple Fusions for Video Denoising 2991
Kai Guo, Seungwon Choi, Jongseong Choi, Lae-Hoon Kim

ThicknessVAE: Learning a Lateral Prior for Clothed Human Body Reconstruction 2996
Xiaotao Wu, Zhaoxin Fan, Huiguang He, Dinggang Shen

Language-Queried Target Sound Extraction Without Parallel Training Data 3001
Hao Ma, Zhiyuan Peng, Xu Li, Yukai Li, Mingjie Shao, Qiuqiang Kong, Ju Liu

CLIPGaze: Zero-Shot Goal-Directed Scanpath Prediction Using CLIP 3006
Yantao Lai, Rong Quan, Dong Liang, Jie Qin

Low-Complexity Neural Speech Dereverberation with Adaptive Target Control 3011
Nagashree K. S. Rao, Srikanth Raj Chetupalli, Shrishti Saha Shetu, Emanuël A. P. Habets, Oliver Thiergart

Interpolation for Weight-Constrained Nested Arrays Having Non-Central ULA Segments in the Coarray 3016
Pranav Kulkarni, P. P. Vaidyanathan

Spatial Frequency Interleaving Residual Autoencoder for Indoor Radio Map Reconstruction 3021
Shitong Chai, Jiahui Li, Mengyao Ma, Junwen Xie, Xiaopeng Fan, Xianqi Zhang

SONIQUE: Video Background Music Generation Using Unpaired Audio-Visual Data 3026
Liqian Zhang, Magdalena Fuentes

KAnoCLIP: Zero-Shot Anomaly Detection Through Knowledge-Driven Prompt Learning and Enhanced Cross-Modal Integration 3031
Chengyuan Li, Suyang Zhou, Jieping Kong, Lei Qi, Hui Xue

PNP-RKD: A Positive-Negative Pair Based Relational Knowledge Distillation Method for Cross-Domain Speaker Verification 3036
Qing Gu, Yan Song, Nan Jiang, Pengfei Cai, Ian McLoughlin

Binary Stochastic Flip Optimization for Training Binary Neural Networks 3041
Tatsukichi Shibuya, Nakamasa Inoue, Rei Kawakami, Ikuro Sato

Linear Time Complexity Conformers with SummaryMixing for Streaming Speech Recognition 3046
Titouan Parcollet, Rogier van Dalen, Shucong Zhang, Sourav Bhattacharya

An End-To-End Graph-Guided Spatiotemporal Model for Adaptive Frame-Level Facial Affect Analysis in the Wild 3051
Yan Liang, Yan Hao, Zenan Yao, Jiacheng Liao, Jiahui Pan

DETCP: Self-Detoxifying Language Models with Contrastive Pairs 3056
Dianqing Liu, Yi Liu, Junbo Guo, Zhendong Mao

Fair MP-BOOST: Fair and Interpretable Minipatch Boosting 3061
Camille Olivia Little, Genevera I. Allen

Ultra Low-Compute Complex Spectral Masking for Multichannel Speech Enhancement 3066
Ashutosh Pandey, Juan Azcarreta

Robust Fixed-Filter Sound Zone Control with Audio-Based Position Tracking 3071
Sankha Subhra Bhattacharjee, Andreas Jonas Fuglsig, Flemming Christensen, Jesper Rindom Jensen, Mads Græsbøll Christensen

Hierarchical Label Propagation: A Model-Size-Dependent Performance Booster for AudioSet Tagging 3076
Ludovic Tuncay, Etienne Labbé, Thomas Pellegrini

3D Gaussian Splatting with Grouped Uncertainty for Unconstrained Images 3081
Hao-Yu Hou, Chia-Chi Hsu, Yu-Chen Huang, Mu-Yi Shen, Wei-Fang Sun, Cheng Sun, Chia-Che Chang, Yu-Lun Liu, Chun-Yi Lee

InfoHarmonizer Graph Contrastive Clustering..... 3086
Zhongyang Zhou, Haomin Wu, Zihao Feng, Feiyu Chen, Bin Tang

CheapNVS: Real-Time On-Device Narrow-Baseline Novel View Synthesis 3091
Konstantinos Georgiadis, Mehmet Kerim Yucel, Albert Saà-Garriga

Low-Resource Text-To-Speech Synthesis Using Noise-Augmented Training of ForwardTacotron 3096
Kishor Kayyar Lakshminarayana, Frank Zalkow, Christian Dittmar, Nicola Pia, Emanuël A.P. Habets

DDNet: Deformable Convolution and Dense FPN for Surface Defect Detection in Recycled Books..... 3101
Jun Yu, WenJian Wang

Sound Zone Control Robust to Sound Speed Change 3106
Sankha Subhra Bhattacharjee, Jesper Rindom Jensen, Mads Græsbøll Christensen

Using Emotionally Rich Speech Segments for Depression Prediction.....3111
Jiawei Yu, Heysem Kaya

TTS-Transducer: End-To-End Speech Synthesis with Neural Transducer.....3116
Vladimir Bataev, Subhankar Ghosh, Vitaly Lavrukhin, Jason Li

Trick-GS: A Balanced Bag of Tricks for Efficient Gaussian Splatting..... 3121
Anil Armagan, Albert Saà-Garriga, Bruno Manganelli, Mateusz Nowak, M. Kerim Yucel

Personalizing Keyword Spotting with Speaker Information 3126
Beltrán Labrador, Pai Zhu, Guanlong Zhao, Angelo Scorza Scarpati, Quan Wang, Alicia Lozano-Diez, Ignacio Lopez-Moreno

Deep Unfolded Approximate Message Passing for Quantitative Acoustic Microscopy Image Reconstruction..... 3131
Odysseas Pappas, Jonathan Mamou, Adrian Basarab, Denis Kouamé, Alin Achim

Multi-Objective Reinforcement Learning for Cognitive Radar Resource Management 3136
Ziyang Lu, Subodh Kalia, M. Cenk Gursoy, Chilukuri K. Mohan, Pramod K. Varshney

Multiband Unlimited Sampling: Super-Nyquist Or Sub-Nyquist, that is the Question! 3141
Ruiming Guo, Gal Shtendel, Ayush Bhandari

Proximity Detection and Trajectory Recognition with Machine Learning for UHF RFID Systems 3146
Thomas M. Pohl, Christoph F. Mecklenbräuker, Holger Arthaber

Psycholinguistic Features Predict Word Duration in Hindi Read Aloud Speech 3150
Rajakrishnan Rajkumar, Sneha Raman, Aadya Ranjan, Mildred Pereira, Nagesh Nayak, Preeti Rao

NCNet: Learning to Find Non-Consistent Correspondence Using Learnable Frequency Response Function 3155
Ruiyuan Li, Zhaolin Xiao, Meng Zhang, Haiyan Jin, Haonan Su

Transitive Inference in Large Language Models and Prompting Intervention 3160
Wenya Wu, Weihong Deng

Event-Driven Prony: Towards Asynchronous Spectral Estimation 3165
Ruiming Guo, Yuliang Zhu, Ayush Bhandari

ARM : nnU-Net with Arena Mechanism for Medical Image Segmentation 3170
Haoran Luo, Cong Guan, Tengfei Shao, Shenglei Li, Tomoji Kishi, Osamu Yoshie

Cooperative ISAC for Localization and Velocity Estimation Using OFDM Waveforms in Cell-Free MIMO Systems 3175
Zihuan Wang, Vincent W.S. Wong

Effective Pre-Training of Audio Transformers for Sound Event Detection 3180
Florian Schmid, Tobias Morocutti, Francesco Foscarin, Jan Schlüter, Paul Primus, Gerhard Widmer

D2-MLP: Dynamic Decomposed MLP Mixer for Medical Image Segmentation 3185
Jin Yang, Jing Yang, Xiaobing Yu, Peijie Qiu, Sunil Prajapat

Estimating Musical Surprisal in Audio 3190
Mathias Rose Bjare, Giorgia Cantisani, Stefan Lattner, Gerhard Widmer

MetaCert: Metabolic Attention Network Utilizing Uncertainty Estimation for Multimodal Aspect-Category-Sentiment Triple Extraction 3195
Haoran Luo, Cong Guan, Tengfei Shao, Shenglei Li, Tomoji Kishi, Osamu Yoshie

Contrast Memory for Unsupervised Anomaly Detection 3200
Jiahao Li, Yiqiang Chen, Yunbing Xing, Yang Gu, Xiangyuan Lan

Dilated Convolution for Time Series Learning 3205
Wang Zhang, Subhro Das, Lam M. Nguyen, Luca Daniel

Hierarchical Relation Distillation for Efficient 3D Visual Grounding 3210
Jun Liu

Faster Speech-LLaMA Inference with Multi-Token Prediction 3215
Desh Raj, Gil Keren, Junteng Jia, Jay Mahadeokar, Ozlem Kalinli

Partially Observable Contextual Bandits with Linear Payoffs 3220
Sihan Zeng, Sujay Bhatt, Alec Koppel, Sumitra Ganesh

Opportunities and Challenges for Bluetooth LE Audio Assistive Listening Systems 3225
Ian C. Bruce, Steve Armstrong, Daniel J. Bosnyak, Hany Tawfik

Improving Multilingual ASR in the Wild Using Simple N-Best Re-Ranking 3230
Brian Yan, Vineel Pratap, Shinji Watanabe, Michael Auli

On Momentum Acceleration for Randomized Coordinate Descent in Matrix Completion 3235
Matthew Callahan, Trung Vu, Raviv Raich

A Multi-Wavelength Optical Sensing Framework for Calibration-Free Wearable Blood Pressure Monitoring 3240
Tarek Hamid, Patricia Flores, Jane Byun, Xi Chen, Haoran Zhang, Kyle Quinn, Amanda Watson

Contextual Speech Extraction: Leveraging Textual History as an Implicit Cue for Target Speech Extraction 3245
Minsu Kim, Rodrigo Mira, Honglie Chen, Stavros Petridis, Maja Pantic

On-Device Hand Model for Robust Gesture Detection 3250
Ivan Bondarets, Veronika Prokhorchuk, Andrii Kozyr, Oleksandr Trunov

Flexible Event-Driven Biological Imaging Via Bayesian Inference 3255
Yayin Tan, Tianle Zhao, Zhiqin Chu

Adaptive Canonical Correlation Analysis with Application to Time Synchronization for Signal Alignment 3260
Spyridon Peppas, Nicholas D. Sidiropoulos

Topology-Informed Pre-Training of Graph Neural Networks 3265
Peiyu Liang, Yulia R. Gel, Yuzhou Chen

Data-Driven Processing Using Parametric Neural Network for Improved Bluetooth Channel Sounding Distance Estimation 3270
Andrii Tsemko, Avik Santra, Oleg Kapshii, Ashutosh Pandey

Enhanced Weakly Supervised Few-Shot Classification & Segmentation 3275
Srinivasa Rao Nandam, Sara Atito, Zhenhua Feng, Josef Kittler, Muhammad Awais

Evaluating the Posterior Sampling Ability of Plug&Play Diffusion Methods in Sparse-View CT 3280
Liam Moroy, Guillaume Bourmaud, Frédéric Champagnat, Jean-François Giovannelli

Low-Shot Image Classification Using Mixture of Experts 3285
Zheng Zhang, Saket Sathe

On the Relation Between Speech Quality and Quantized Latent Representations of Neural Codecs 3290
Mhd Modar Halimeh, Matteo Torcoli, Philipp Grundhuber, Emanuël A. P. Habets

Robust Audio Deepfake Detection Using Ensemble Confidence Calibration 3295
Chin Yuen Kwok, Duc-Tuan Truong, Jia Qi Yip

Extending Whisper for Emotion Prediction Using Word-Level Pseudo Labels 3300
Chin Yuen Kwok, Sheng Li, Jia Qi Yip, Chenhui Chu, Tatsuya Kawahara, Eng Siong Chng

Joint Semantic Knowledge Distillation and Masked Acoustic Modeling for Full-Band Speech Restoration with Improved Intelligibility 3305
Xiaoyu Liu, Xu Li, Joan Serrà, Santiago Pascual

Speech-N-LlaMA: Improving Speech LLMs with Multi-Pass Training 3310
Amit Kumar Singh Yadav, Gil Keren, Desh Raj, Wei Zhou, Junteng Jia, Ke Li, Ying Xu, Chunyang Wu, Jay Mahadeokar, Ozlem Kalinli

Generative Speech Foundation Model Pretraining for High-Quality Speech Extraction and Restoration 3315
Pin-Jui Ku, Alexander H. Liu, Roman Korostik, Sung-Feng Huang, Szu-Wei Fu, Ante Jukić

Rethinking Mamba in Speech Processing by Self-Supervised Models 3320
Xiangyu Zhang, Jianbo Ma, Mostafa Shahin, Beena Ahmed, Julien Epps

Learning Class Unique Features in Fine-Grained Visual Classification 3325
Runkai Zheng, Li Liu, Zhijia Yu, Yinqi Zhang, Hei Victor Cheng, Chris Ding

Quantum Reinforcement Learning for Coordinated Satellite Systems 3330
Gyu Seon Kim, Samuel Yen-Chi Chen, Soohyun Park, Joongheon Kim

Hierarchical Context Interaction and Reasoning with Transformer for Emotion Recognition 3335
Wenxuan Wang, Chenglei Wang, Xiaomei Wang, Xuli Shen, Qing Xu, Xuelin Qian

Multi-Agent Decision Transformer for Power Control in Wireless Networks 3340
Yiming Zhang, Kun Yang, Cong Shen, Dongning Guo

Cramér-Rao Bounds for Wideband Near-Field Sensing 3345
Tong Wei, Kumar Vijay Mishra, Linlong Wu, Bhavani Shankar Mysore Rama Rao

Robust and Efficient Adversarial Defense in SNNs Via Image Purification and Joint Detection 3350
Weiran Chen, Qi Xu

SAM-OCTA2: Layer Sequence OCTA Segmentation with Fine-Tuned Segment Anything Model 2 3355
Xinrun Chen, Chengliang Wang, Haojian Ning, Mengzhan Zhang, Mei Shen, Shiying Li

Convolutional Prompting for Broad-Domain Retinal Vessel Segmentation 3360
Qijie Wei, Weihong Yu, Xirong Li

DRDM: A Disentangled Representations Diffusion Model for Synthesizing Realistic Person Images 3365
Enbo Huang, Yuan Zhang, Faliang Huang, Guangyu Zhang, Yang Liu

ReCLAP: Improving Zero Shot Audio Classification by Describing Sounds 3370
Sreyan Ghosh, Sonal Kumar, Chandra Kiran Reddy Evuru, Oriol Nieto, Ramani Duraiswami, Dinesh Manocha

Precision in Pathology: PMA-DETR Elevates Tumor Lesion Detection 3375
Yecheng Zhao, ZiHan Zhou, Lei Qi, Hui Xue

Overcoming Uncertain Incompleteness for Robust Multimodal Sequential Diagnosis Prediction Via Curriculum Data Erasing Guided Knowledge Distillation 3380
Heejoon Koo

Joint Task Offloading and Routing in Wireless Multi-Hop Networks Using Biased Backpressure Algorithm 3385
Zhongyuan Zhao, Jake Perazzone, Gunjan Verma, Kevin Chan, Ananthram Swami, Santiago Segarra

Maximum Likelihood Estimation of Stable ARX Models Using Randomized Coordinate Descent 3390
Ozmen Erkin Kokten, Raviv Raich

Semantic-Aware Prompt Learning for Multimodal Sarcasm Detection 3395
Guangjin Wang, Bao Wang, Fuyong Xu, Zhenfang Zhu, Peipei Wang, Ru Wang, Peiyu Liu

High-Fidelity Single-View Reconstruction of Indoor Scenes Using 3D Shape Prior Template and Pixel-Aligned Deformation 3400
Xiaohao Zhang, Xiaolin He, Jialin Wu, Xu Wang, Zhuo Tang, Ruihui Li

Multi-Stage Multimodal Distillation for Audio-Visual Speaker Tracking 3405
Yidi Li, Wenkai Zhao, Zeyu Wang, Zhenhuan Xu, Bin Ren, Nicu Sebe

Node-Centric Meta Structure Search in Heterogeneous Graphs 3410
Xiaoou Zhang, Yang Gao, Yang Liu, Yujia Zhu, Chuan Zhou, Peng Zhang, Yong Ding, Qingyun Liu, Hongyang Chen

Data Glove-Based Personalized Continuous Gesture Segmentation 3415
Liufeng Fan, Zhan Zhang, Yiwei Wang, Decheng Zuo, Yinran Wang, Zhongyuan Chen

A Multi-Scenario Attention-Based Generative Model for Personalized Blood Pressure Time Series Forecasting 3420
Cheng Wan, Chenjie Xie, Longfei Liu, Dan Wu, Ye Li

Watermarking Datasets for LLM Fine-Tuning 3425
Jing Qiu, Xi Yang, Shuai Li, Kejiang Chen, Weiming Zhang, Nenghai Yu

AdaptiveDrop: A Simple Adaptive Label Noise Filtering Scheme for Enhanced Self-Supervised Speaker Verification 3430
Abderrahim Fathan, Xiaolin Zhu, Jahangir Alam

Towards Personalized Federated Learning Via Contrastive-Augmented Local Memorization Retrieval 3435
Peifeng Zhang, Jiahui Chen

Improving Compressive Imaging Recovery Via Measurement Augmentation 3440
Romario Gualdrón-Hurtado, Roman Jacome, Leon Suarez, Emmanuel Martinez, Henry Arguello

Reinforcement Learning-Based Multi-Teacher Knowledge Distillation for Enhancing Retrieval Ranking Consistency 3445
Xiukang Yang, Jingguo Ge, Liangxiong Li, Bingzhen Wu

Efficient Anchor Graph Clustering Through Enhanced Within-Cluster Homogeneity 3450
Fangyuan Xie, Lin Zhao, Jingjing Xue, Feiping Nie, Weizhong Yu, Xuelong Li

Advancing SAR Image Robustness: Integrating Diffusion Models for Adversarial Purification and Speckle Noise Suppression 3455
Ganglin Xie, Haobo Lu, Xinze Zhang, Kun He

CIEGCL: Counterfactual Intervention Enhancing Graph Contrastive Learning in Implicit Feedback 3460
Zijian Huang, Tian Jiang, Yan Feng, Zerui Wen, Xiaohui Cui

Non-Autoregressive Multimodal Machine Translation 3465
Guojing Liu, Xiangqian Ding, Huili Gong, Xiangyu Qu, Zhenyu Yang, Kai Yan

Learning to Reconstruct Signals with Inexact Sensing Operator Via Knowledge Distillation 3470
Roman Jacome, Leon Suarez, Romario Gualdrón-Hurtado, Luis Gonzalez, Henry Arguello

Transfer Learning with Transformer and LSTM for Digital Pre-Distortion of Terahertz/mmWave Transceiver 3475
Gouheng Zhao, Kai Ying, Qingsong Wen, Junwen Zhang, Lin Gui

A High-Precision Character Cartoon Style Transfer Method Based on VToonify and Diffusion Models 3480
Weiting Wang, Weiqi Wang, Feilong Bao

Mamba for Streaming ASR Combined with Unimodal Aggregation 3485
Ying Fang, Xiaofei Li

Crowdsourced Homophily Ties Based Graph Annotation Via Large Language Model 3490
Yu Bu, Yulin Zhu, Kai Zhou

Gradient Norm-Based Fine-Tuning for Backdoor Defense in Automatic Speech Recognition 3495
Nanjun Zhou, Weilin Lin, Li Liu

Knowledge is Powerful: Art Knowledge-Driven Framework for Painting Style Classification Integrating Multimodal Knowledge 3500
Haoyang Chen, Chongjun Wang, Yi Xin, Lei Zhang

Continuously Learning New Words in Automatic Speech Recognition 3505
Christian Huber, Alexander Waibel

State-Augmented Opportunistic Routing in Wireless Communication Systems with Graph Neural Networks 3510
Sourajit Das, Navid NaderiAlizadeh, Alejandro Ribeiro

Automatic Text Pronunciation Correlation Generation and Application for Contextual Biasing 3515
Gaofeng Cheng, Haitian Lu, Chengxu Yang, Xuyang Wang, Ta Li, Yonghong Yan

Description-Based Controllable Text-To-Speech with Cross-Lingual Voice Control 3520
Ryuichi Yamamoto, Yuma Shirahata, Masaya Kawamura, Kentaro Tachibana

SLIDE: Integrating Speech Language Model with LLM for Spontaneous Spoken Dialogue Generation 3525
Haitian Lu, Gaofeng Cheng, Liuping Luo, Leying Zhang, Yanmin Qian, Pengyuan Zhang

Beyond Speaker Identity: Text Guided Target Speech Extraction 3530
Mingyue Huo, Abhinav Jain, Cong Phuoc Huynh, Fanjie Kong, Pichao Wang, Zhu Liu, Vimal Bhat

Speech Recognition Rescoring with Large Speech-Text Foundation Models 3535
Prashanth Gurunath Shivakumar, Jari Kolehmainen, Aditya Gourav, Yi Gu, Ankur Gandhe, Ariya Rastrow, Ivan Bulyko

Group-CLIP Uncertainty Modeling for Group Re-Identification 3540
Qingxin Zhang, Haoyan Wei, Yang Qian

A Near-Field 3D Parameter Estimation Method Based on a Symmetric Enhanced Nested Array 3545
Linke Yu, Hua Chen, Dingfan Xue, Wei Liu, Ye Tian, Gang Wang

MotionComposer: Enhancing Rhythmic Music Generation with Adaptive Retrieval Reference 3550
Jinting Wang, Li Liu, Jun Wang

Fine-Portraitist: Visualizing the Speaker's Face Portrait During Speech Listening 3555
Jinting Wang, Li Liu, Jun Wang

SiQA: A Large Multi-Modal Question Answering Model for Structured Images Based on RAG 3560
Jiawang Liu, Ye Tao, Fei Wang, Hui Li, Xiugong Qin

RWKVMatch: Vision RWKV-Based Multi-Scale Feature Matching Network for Unsupervised Deformable Medical Image Registration 3565
Zixuan He, Jing Tang, Zitong Zhao, Zeyu Gong

Low-Rate Modulo Folded ADC for Detecting Linearly Modulated Communication Symbols 3570
Satish Mulleti, Kumar Appaiah, Sibi Raj B. Pillai

Continual Self-Supervised Learning Considering Medical Domain Knowledge in Chest CT Images 3575
Ren Tasai, Guang Li, Ren Togo, Minghui Tang, Takaaki Yoshimura, Hiroyuki Sugimori, Kenji Hirata, Takahiro Ogawa, Kohsuke Kudo, Miki Haseyama

Combining Loss-Aware Curriculum Learning with Incomplete Graph Neural Networks 3580
Jiawei Chen, Keao Xi, Gaoke Zhang, Yueheng Sun, Wenjun Wang

A Lowrate Variable-Bias Integrate-And-Fire Time Encoding Machine 3585
Anshu Arora, Satish Mulleti

Rethinking Early-Fusion Strategies for Improved Multimodal Image Segmentation 3590
Zhengwen Shen, Yulian Li, Han Zhang, Yuchen Weng, Jun Wang

Generative Dataset Distillation Based on Self-Knowledge Distillation 3595
Longzhen Li, Guang Li, Ren Togo, Keisuke Maeda, Takahiro Ogawa, Miki Haseyama

DeformAvatar: Point-Based Human Avatar Re-Targeting and Rendering 3600
Renyi Zhan, Zhi-Qi Cheng, Junyao Chen, Xiaojiang Peng

Automatic Parkinson's Disease Detection from Speech: Layer Selection Vs Adaptation of Foundation Models 3605
Tilak Purohit, Barbara Ruvolo, Juan Rafael Orozco-Arroyave, Mathew Magimai.-Doss

Emotion Information Recovery Potential of Wav2vec2 Network Fine-Tuned for Speech Recognition Task 3610
Tilak Purohit, Mathew Magimai-Doss

Enhancing Time Series Prediction with Evolutionary Algorithm-Based Optimization of LSTM 3615
Jingyu Sun, Hanting Zhang, Jianfeng Wang

Identical-Delay Based 2-D DOA and Frequency Joint Estimation with Sub-Nyquist Sampling for URA 3620
Liang Liu, Xinyun Zhang, Xinyi Zhou, Lu Gan, Jiancheng An, Hongbin Li

DCCMamba: A Dual-Stream Cross-Time and Cross-Feature with Mamba for Multivariate Time Series Forecasting 3625
Senlin Liang, Zhuoyue Wang, Chunna Zhao, Yaqun Huang, Jinpeng Xu, Yaoyuan Yang

A New Model for Prototype-Based Continual Learning in Hyperspherical Space 3630
Yixin Ren, Yewei Xia, Li Ke, Longtao Huang, Hui Xue, Shuigeng Zhou

A Two-Timescale Primal-Dual Algorithm for Decentralized Optimization with Compression 3635
Haoming Liu, Chung-Yiu Yau, Hoi-To Wai

NCL-CIR: Noise-Aware Contrastive Learning for Composed Image Retrieval 3640
Peng Gao, Yujian Lee, Zailong Chen, Xubo Liu, Hui Zhang, Yiyang Hu, Guquan Jing

A Task-Oriented Real-Time and Robust Feature Compression and Selection Method in Collaborative Intelligence System 3645
Kaile Wang, Tantan Zhao, Fan Li

Enhancing Document-Level Relation Extraction Through Entity-Pair-Level Interaction Modeling 3650
Wanlong Liu, Dingyi Zeng, Li Zhou, Yichen Xiao, Malu Zhang, Wenyu Chen

Towards Patronizing and Condescending Language in Chinese Videos: A Multimodal Dataset and Detector 3655
Hongbo Wang, Junyu Lu, Yan Han, Kai Ma, Liang Yang, Hongfei Lin

De-Confusing Hard Samples for Text Semantic Hashing 3660
Tian Huang, Jian Wang, Yuqing Sun

Boosting Jailbreak Attack with Momentum 3665
Yihao Zhang, Zeming Wei

UMETTS: A Unified Framework for Emotional Text-To-Speech Synthesis with Multimodal Prompts..... 3670
Xiang Li, Zhi-Qi Cheng, Jun-Yan He, Junyao Chen, Xiaomao Fan, Xiaojiang Peng, Alexander G. Hauptmann

Graph Refinement in Latent Space: A Hypergraph Convolution for Underwater Object Detection 3675
Meghna Kapoor, Badri Narayan Subudhi, Ankur Bansal

DMKPN: Image Deblurring Under Multi-Factor Aliasing Diffusion Degradation..... 3680
Ying Zhang, Xiongxin Tang, Hanxiang Yang, Qiao Chen, Fanjiang Xu

Self-Convolutional Attention-Based Uncertainty-Aware Network for Single-Image Super-Resolution..... 3685
Jinbin Wang, Aiping Yang, Zihao Wei, Qinghua Hu

Multi-User Non-Orthogonal Multiple Access in Power Line Communications Under Generalized Gaussian Noise 3690
Roopesh Ramesh, Sanjeev Gurugopinath, R. Muralishankar

Impact of Glyph Information on Latent Space Diffusion Models for Accurate Handwritten Text Generation 3695
Ying-Li Lin, Hao-Chung Cheng, Chung-I Huang, Chien-Yao Wang, Jia-Ching Wang

CT Image Prediction of PD-1 Gastric Cancer Patients Based on the PLSG Framework 3700
Chaoyu Yuan, Nan Wang, Mohan Wang, Yingwei Xue

PEPL: Precision-Enhanced Pseudo-Labeling for Fine-Grained Image Classification in Semi-Supervised Learning..... 3705
Bowen Tian, Songning Lai, Lujundong Li, Zhihao Shuai, Runwei Guan, Tian Wu, Yutao Yue

Cluster-Refined Optimal Transport for Unsupervised Action Segmentation..... 3710
Shijie Wang, Jinrong Zhang, Yule Liu, Shiyao Li, Lin Feng

Disparity-Guided Cross-View Transformer for Stereo Image Super-Resolution 3715
Bingting Li, Wenjing Shang, Yongshun Gong, Qiangchang Wang, Xinxin Zhang, Yilong Yin

DeBeauty: A Joint Framework for Facial Beautification Removal Based on Spatial Collaborative Adaptation and Hyperplane Relocation..... 3720
Jinghang Wang, Yi Li, Zexing Zhang, Cunrui Zou, Zibo Liu, Zhiguo Zhou

Sentiment Analysis for Live Video Comments with Diffused Fusion of Cross-Modal Representations 3725
Changfan Luo, Ling Fang, Bensheng Qiu

GDDA: Semantic OOD Detection on Graphs Under Covariate Shift Via Score-Based Diffusion Models 3730
Zhixia He, Chen Zhao, Minglai Shao, Yujie Lin, Dong Li, Qin Tian

Multimodal and Multiple Prompts for Biology 3735
Chonghuinan Wang, Xi Chen, Hongxun Yao

Active Learning for Long-Tailed Annotation 3740
Lin Geng, Ningzhong Liu, Han Sun, Jie Qin

Simple Adaptive Spectrum Graph Filters for Rumor Detection 3745
Nanjun Yu, Qiang Cao, Zheng Dong

CMGait: Enhancing Cross-Modality Gait Recognition Between LiDAR and RGB Through Contrastive Identity-Consistent Feature Aggregation 3750
Yubo Wang, Bin Liu, Zhiwei Zhao, Jixiang Niu, Qi Chu, Nenghai Yu

Symmetry and Fusion Data Augmentation for Semi-Supervised Medical Segmentation 3755
Yishan Zhang, Wenxin Yu, Zhiqiang Zhang, Jun Gong, Peng Chen, Chang Liu

Time-Independent Spiking Neuron Via Membrane Potential Estimation for Efficient Spiking Neural Networks 3760
Hanqi Chen, Lixing Yu, Shaojie Zhan, Penghui Yao, Jiankun Shao

STAD: Joint Spatial-Temporal Dimension and Channel Correlation for Time Series Anomaly Detection 3765
Mingle Zhou, Xingli Wang, Delong Han, Jin Wan, Jiachen Li

Community-Entropy Based Graph Structure Learning for Topology-Imbalance 3770
Xinyi Wang, Ling Guo, Hui Yan

Camouflaged Object Detection with CNN-Transformer Harmonization and Calibration 3775
Yilin Zhao, Qing Zhang, Yuetong Li

Tag-Aware Weakly-Supervised Online Hashing with Enhanced Joint Representation 3780
Na Wang, Yu-Wei Zhan, Zhen-Duo Chen, Yongxin Wang, Xin Luo, Xin-Shun Xu

Dual-Energy CT Metal Artifact Reduction by Combined Material Decomposition and Projection Domain Threshold Segmentation 3785
Kai Chen, Tianling Lyu, Jean-Louis Coatrieux, Yang Chen

The Potential of Speech Features to Discriminate Between Original and Machine-Translated Texts 3790
Yongjian Chen, Mireia Farrús, Antonio Toral

MTTM: Memory-Augmented with Mamba for 3D Medical Images Analysis 3795
Hongkai Wei, Yang Yang, Shijie Sun, Huansheng Song, Keyu Guo, Yongfeng Bu

Optimized Dynamic Watermarking for Audio DNNs with Adaptive Embedding and Boundary Sampling 3800
Hao Fei, Hewang Nie, Siqi Sun, Songfeng Lu, Ting Luo, Ling Qian, Dunbo Cai, Zhiguo Huang, Runqing Zhang

HandS3C: 3D Hand Mesh Reconstruction with State Space Spatial Channel Attention from RGB Images 3805
Zixun Jiao, Xihan Wang, Zhaoqiang Xia, Lianhe Shao, Quanli Gao

3DGCQA: A Quality Assessment Database for 3D AI-Generated Contents 3810
Yingjie Zhou, Zicheng Zhang, Farong Wen, Jun Jia, Yanwei Jiang, Xiaohong Liu, Xiongkuo Min, Guangtao Zhai

Improved Feature Extraction Network for Neuro-Oriented Target Speaker Extraction 3815
Cunhang Fan, Youdian Gao, Zexu Pan, Jingjing Zhang, Hongyu Zhang, Jie Zhang, Zhao Lv

WeightedKV: Attention Scores Weighted Key-Value Cache Merging for Large Language Models 3820
Jian Yuan, Ziwei He, Haoli Bai, Jingwen Leng, Bo Jiang

Padnet: A Patch-Based Anomaly Detection Framework for Industrial Pipeline Damage Detection 3825
Erofili Alexaki, Christos Papaioannidis, Vasileios Mygdalis, Ioannis Pitas

Complementary Learning System Theory-Based Active Learning for Audio Classification 3830
Hui Geng, Zijian Gao, Tianjiao Wan, Dawei Feng, Changjian Wang, Kele Xu

PB-UAP: Hybride Universal Adversarial Attack for Image Segmentation 3835
Yufei Song, Ziqi Zhou, Minghui Li, Xianlong Wang, Hangtao Zhang, Menghao Deng, Wei Wan, Shengshan Hu, Leo Yu Zhang

LV-ReID: Large Language-Vision Alignment Model for Text-Based Person Re-Identification 3840
Yinghui Xia, Chao Wang, Jinsong Yang

Enhancing Emotion Recognition in Incomplete Data: A Novel Cross-Modal Alignment, Reconstruction, and Refinement Framework 3845
Haoqin Sun, Shiwan Zhao, Shaokai Li, Xiangyu Kong, Xuechen Wang, Jiaming Zhou, Aobo Kong, Yong Chen, Wenjia Zeng, Yong Qin

Adapting Whisper for Code-Switching Through Encoding Refining and Language-Aware Decoding 3850
Jiahui Zhao, Hao Shi, Chenrui Cui, Tianrui Wang, Hexin Liu, Zhaoheng Ni, Lingxuan Ye, Longbiao Wang

Vector Quantized Diffusion Model Based Speech Bandwidth Extension 3855
Yuan Fang, Jinglin Bai, Jiajie Wang, Xueliang Zhang

MSTBI: Head CT Detection and Prognostic Assessment of Traumatic Brain Injury Dataset 3860
Yang Xu, Menghao Fang, Qiuyu Fu, Mengqi Qu, Junyao Chen, Zexian Xie, Shike Hou, Lu Lu

GraphDAE-PU: Graph Denosing Auto-Encoder for Arbitrary-Scale Point Cloud Upsampling 3865
Yuzhong Deng, Zhiheng Su, Penghui Shang, Dongzhen Liu, Di Wu, Jianxiao Zou, Shicai Fan

Spiking Generative Models Based on Variational Autoencoder and Adversarial Training 3870
Wenchuan Zhang, Jie Zhang, Ricky Yuen-Tan Hou, Weifeng Su, Wentao Fan

Target Speaker ASR with Whisper 3875
Alexander Polok, Dominik Klement, Matthew Wiesner, Sanjeev Khudanpur, Jan Černocký, Lukáš Burget

Stochastic-Aware Mamba Diffusion for Pedestrian Trajectory Prediction 3880
Ziyang Ren, Ping Wei, Haowen Tang, Huan Li, Jin Yang, Jialu Qin

Optimal Combination Policies for Error Exponent Maximization in Social Learning 3885
Ping Hu, Mert Kayaalp, Ali H. Sayed

Label-Constrained Unsupervised Domain Adaptation for Semantic Segmentation with Diffusion Models 3890
Alexandre Stenger, Étienne Baudrier, Benoît Naegel, Nicolas Passat

MSEMG: Surface Electromyography Denoising with a Mamba-Based Efficient Network 3895
Yu-Tung Liu, Kuan-Chen Wang, Rong Chao, Sabato Marco Siniscalchi, Ping-Cheng Yeh, Yu Tsao

MADiff: Text-Guided Fashion Image Editing with Mask Prediction and Attention-Enhanced Diffusion 3900
Zechao Zhan, Dehong Gao, Jinxia Zhang, Jiale Huang, Yang Hu, Xin Wang

CroPrompt: Cross-Task Interactive Prompting for Zero-Shot Spoken Language Understanding 3905
Libo Qin, Fuxuan Wei, Qiguang Chen, Jingxuan Zhou, Shijue Huang, Jiasheng Si, Wenpeng Lu, Wanxiang Che

Enhancing Zero-Shot Cross-Lingual Event Argument Extraction with Language-Independent Information 3910
Xiruijie Yi, Xiaoxu Zhu, Peifeng Li

Decreasing Word Error Rates in Paragraph Handwritten Text Recognition with Synthetic Data 3915
Ernest Yu-Kai Chew, Adams Wai-Kin Kong, Lim Joo-Hwee

Cross-Attention Inspired Selective State Space Models for Target Sound Extraction 3920
Donghang Wu, Yiwen Wang, Xihong Wu, Tianshu Qu

SymGaussian: Occluded Human Rendering with Multi-Scale Symmetry Feature from Monocular Video 3925
Zekai Jiang, Tong Duan, Dongyu Zhang

LLaVA-SG: Leveraging Scene Graphs as Visual Semantic Expression in Vision-Language Models 3930
Jingyi Wang, Jianzhong Ju, Jian Luan, Zhidong Deng

Binary Representation Learning for Discriminative Acoustic Unit Discovery 3935
Rui Niu, Jie Chen, Long Ma, Changhe Song, Weihao Wu, Zhiyong Wu

Optimizing Dual-Mode UAV-Assisted Remote IoT Data Collection with Decentralized DRL in 6G-Enabled Space-Air-Ground Networks 3940
Ru Jin, Rongheng Lin

Learning Primitive Relations for Compositional Zero-Shot Learning 3945
Insu Lee, Jiseob Kim, Kyuhong Shim, Byonghyo Shim

TA-V2A: Textually Assisted Video-To-Audio Generation 3950
Yuhuan You, Xihong Wu, Tianshu Qu

SNNPTrack: Spiking Neural Network Based Prompt for High-Accuracy RGBE Tracking 3955
Yixi Ji, Qinghang Zhao, Yuping Liang, Jinjian Wu

CAF-YOLO: A Robust Framework for Multi-Scale Lesion Detection in Biomedical Imagery 3960
Zilin Chen, Shengnan Lu

Self-Supervised Learning and Image-Prompt Fusion for AIGC Image Quality Assessment 3965
Yan Zhao, Qingbing Sang, Zhaohong Deng, Xiaojun Wu

ImageFlowNet: Forecasting Multiscale Image-Level Trajectories of Disease Progression with Irregularly-Sampled Longitudinal Medical Images 3970
Chen Liu, Ke Xu, Liangbo L. Shen, Guillaume Huguet, Zilong Wang, Alexander Tong, Danilo Bzdok, Jay Stewart, Jay C. Wang, Lucian V. Del Priore, Smita Krishnaswamy

Exploring Generalization Boundaries of Unsupervised Industrial Anomaly Detection Models Through Attribute Perturbation 3975
Wei Ran, Yuzhuo Fu, Ting Liu

Long-Range Multi-Scale Fusion for Efficient Single Image Super-Resolution 3980
Yaohua Li, Xiaole Zhao, Xiaobo Zhang, Yan Yang, Tianrui Li

Importance-Awareness Masking Network for Robust Document Retrieval 3985
Junping Liu, Jiaqi He, Xinrong Hu, Wangli Yang, Jie Yang, Yi Guo

Harmonizing for Defect Visibility with Fine-Grained Hierarchical Interaction Learning 3990
Zhongze Wu, Yitian Long, Xiu Su, Yueyi Luo, Shan You, Jun Long

PET: High-Frequency Temporal Self-Consistency Learning for Partially Deepfake Audio Localization 3995
Jiayi He, Jiangyan Yi, Jianhua Tao, Siding Zeng

Double Domain Converter Transformer for Improving EEG-Based Emotion Recognition from Video to Game Scenarios 4000
Jun-Yu Pan, Hao-Long Yin, Wei-Long Zheng

DMIBot: Dynamic Multimodal Interaction for Twitter Bot Detection 4005
Xiezhuo Lin, Qingfeng Wu, Yurui Huang

Structural-Aware Disentangled Learning with CLIP for Hyperbolic Zero-Shot Sketch-Based Image Retrieval* 4010
Qing Zhang, Jing Zhang, Feilong Bao, Xiangdong Su, Guanglai Gao

Adversarial Speech-Text Pre-Training for Speech Translation 4015
Chenxuan Liu, Liping Chen, Weitai Zhang, Xiaoxi Li, Peiwang Tang, Mingjia Yu, Sreyan Ghosh, Zhongyi Ye

SVTNet: Dual Branch of Swin Transformer and Vision Transformer for Monocular Depth Estimation 4020
Shuwen Jia, Yongxiong Wang, Han Chen, Shuai Huang

Stable-TTS: Stable Speaker-Adaptive Text-To-Speech Synthesis Via Prosody Prompting 4025
Wooseok Han, Minki Kang, Changhun Kim, Eunho Yang

A Unified Spatiotemporal Frequency Graph Neural Network for fMRI-Based Brain Functional Connectivity Analysis 4030
Yulang Huang, Zhiyuan Ding, Guokai Duan, Yan Liu, Xiangzhu Zeng, Zheng Wang, Yingying Xu, Ling Wang

Modality-Invariant Bidirectional Temporal Representation Distillation Network for Missing Multimodal Sentiment Analysis 4035
Xincheng Wang, Liejun Wang, Yinfeng Yu, Xinxin Jiao

FreqSense: Universal and Low-Latency Adversarial Example Detection for Speaker Recognition with Interpretability in Frequency Domain 4040
Yihuan Huang, Yuanzhe Li, Yanzhen Ren, Weiping Tu, Yuhong Yang

High-Fidelity Music Vocoder Using Neural Audio Codecs 4045
Luca A. Lanzendörfer, Florian Grötschla, Michael Ungersböck, Roger Wattenhofer

LLaQo: Towards a Query-Based Coach in Expressive Music Performance Assessment 4050
Huan Zhang, Vincent K.M. Cheung, Hayato Nishioka, Simon Dixon, Shinichi Furuya

Scaling a Simple Approach to Zero-Shot Speech Recognition 4055
Jinming Zhao, Vineel Pratap, Michael Auli

SKE-MSA: Enhancing Representation Learning with VAD Lexicon for Multimodal Sentiment Analysis 4060
Yang Yang, Xunde Dong, Yupeng Qiang, Wenjie Si

PRimuS: Pretraining IMU Encoders with Multimodal Self-Supervision 4065
Arnav M. Das, Chi Ian Tang, Fahim Kawsar, Mohammad Malekzadeh

Advancing Single-Snapshot DOA Estimation with Siamese Neural Networks for Sparse Linear Arrays 4070
Ruxin Zheng, Shunqiao Sun, Hongshan Liu, Yimin D. Zhang

A Novel Self-Supervised Contrastive Learning Framework for Masked EEG Motor Imagery Modeling 4075
Kunkun Zhang, Qianwei Zhou, Haigen Hu

When CLIP Meets PHOC: A Dual-Branch Network for Historical Document Image Retrieval 4080
Jing Zhang, Hongxi Wei, Qing Zhang

A Spherical-Harmonic Domain Selective Spatial Active Noise Control System Based on Sound Field Reproduction 4085
Huawei Zhang, Huiyuan June Sun, Jihui Aimee Zhang, Prasanga Samarasinghe, Yile Angela Zhang

CTGDiff: A Conditional Diffusion Model for Cardiotocography Signal Synthesis 4090
Xiaoqing Li, Pufan Cai, Yu Lu, Shijie Shi, Liangkun Ma, Xianghua Fu

Dynamic Frequency-Adaptive Knowledge Distillation for Speech Enhancement 4095
Xihao Yuan, Siqi Liu, Hanting Chen, Lu Zhou, Jian Li, Jie Hu

Enhanced Multimodal Emotion Recognition in Conversations Via Contextual Filtering and Multi-Frequency Graph Propagation 4100
Huan Zhao, Yingxue Gao, Haijiao Chen, Bo Li, Guanghui Ye, Zixing Zhang

SSFSL: Self-Supervised and Few-Shot Learning for Cross-Domain Hyperspectral Image Classification 4105
Guohua Lv, Xiang Gao, Qiang Chi, Guixin Zhao, Aimei Dong, Wei Li

Face-StyleSpeech: Enhancing Zero-Shot Speech Synthesis from Face Images with Improved Face-To-Speech Mapping 4110
Minki Kang, Wooseok Han, Eunho Yang

No Data Required: Zero-Shot Domain Adaptation for Automatic Music Transcription 4115
Andrew McLeod

BrainChat: Interactive Semantic Information Decoding from fMRI Using Large-Scale Vision-Language Pretrained Models 4120
Wanqiu Huang, Ke Ma, Tingyu Xie, Hongwei Wang

Fast and Robust High Resolution Frequency Estimation of Damped Signals 4125
Qi Dai, Ruiming Guo, Thierry Blu

Biodenoising: Animal Vocalization Denoising Without Access to Clean Data 4130
Marius Miron, Sara Keen, Jen-Yu Liu, Benjamin Hoffman, Masato Hagiwara, Olivier Pietquin, Felix Effenberger, Maddie Cusimano

EEG Decoding and Visual Reconstruction Via 3D Geometric with Nonstationarity Modelling 4135
Xin Xiao, Kaiwen Wei, Jiang Zhong, Xuekai Wei, Jielu Yan

GAN-Based Speech Enhancement for Low SNR Using Latent Feature Conditioning 4140
Shrishti Saha Shetu, Emanuël A. P. Habets, Andreas Brendel

GPT-C: Generative PrompT Compression 4145
Lijun Liu, Rui Wang, Lihua Jing, Feixiao Lv, Zixuan Zhu

Uncertainty-Aware Correspondence Distillation for Deep Image Clustering 4150
Luyao Chang, Leiting Chen, Chuan Zhou

Semantic Residual for Multimodal Unified Discrete Representation 4155
Hai Huang, Shulei Wang, Yan Xia

An Underwater Image Quality Dataset with Renewed Pairwise Voting 4160
Mengjiao Shen, Miao Yang, Hansen Zhang, Jinyang Zhong, Yuquan Qiu, Jinwei Gu

LogSI: A Benchmark for System-Incremental Log Analysis 4165
Mingjie Zhou, Weidong Yang, Lipeng Ma, Sihang Jiang, Bo Xu, Yanghua Xiao

DSINet: Towards Real-Time Target Speaker Extraction with Dynamic Speaker Information Fusion 4170
Fengyuan Hao, Andong Li, Xiaodong Li, Chengshi Zheng

Prototype Based Masked Audio Model for Self-Supervised Learning of Sound Event Detection 4175
Pengfei Cai, Yan Song, Nan Jiang, Qing Gu, Ian McLoughlin

Hierarchical Prompt Tuning for System-Incremental Log Analysis 4180
Mingjie Zhou, Weidong Yang, Lipeng Ma, Sihang Jiang, Bo Xu, Yanghua Xiao

SDAFE: A Dual-Filter Stable Diffusion Data Augmentation Method for Facial Expression Recognition 4185
Minghao Zhao, Yifei Chen, Jiahao Lyu, Shuangli Du, Zhiyong Lv, Lin Wang

A Mamba-Based Network for Semi-Supervised Singing Melody Extraction Using Confidence Binary Regularization 4190
Xiaoliang He, Kangjie Dong, Jingkai Cao, Shuai Yu, Wei Li, Yi Yu

Enhancing Continual Learning for Medical Imaging: Efficient Knowledge Transfer and Multi-Disease Prediction 4195
Enzhi Wang, Qicheng Li, Di Liu, Bo Yang

FlashSR: One-Step Versatile Audio Super-Resolution Via Diffusion Distillation 4200
Jaekwon Im, Juhan Nam

KCGAFormer: When Large-Kernel ConvFormer Meets KAN in Semantic Segmentation 4205
Shihao Wang, Zhengxing Huang, Enguang Zuo, Alimjan Aysa, Kurban Ubul

STE-Mamba: Automated Multimodal Depression Detection Through Emotional Analysis and Spatio-Temporal Information Ensemble 4210
Zulong Lin, Yaowei Wang, Yujue Zhou, Fei Du, Yun Yang

Group Modeling and Recommendation Based on Multi-Behavior Interactions in Live Streaming E-Commerce 4215
Li Yu, Shuyang Sheng, Yonggang Liu, Yilin Wei, Junyao Xiang

Leveraging Self-Supervised Learning for Speaker Diarization 4220
Jiangyu Han, Federico Landini, Johan Rohdin, Anna Silnova, Mireia Diez, Lukáš Burget

InstAD: Instance-Aware Segmentation Framework for Zero-Shot Multi-Instance Anomaly Detection 4225
Cheng-Yu Ho, Shang-Hong Lai

FCConDubber: Fine and Coarse Grained Prosody Alignment for Expressive Video Dubbing Via Contrastive Audio-Motion Pretraining 4230
Qiulin Li, Zhichao Wu, Hanwei Li, Xin Dong, Qun Yang

Neural Adaptive Contextual Video Streaming 4235
Jianchao Yang, Mufan Liu, Puyue Hou, Yiling Xu, Jun Sun

Evaluation of Wearable Head BCG for PTT Measurement in Blood Pressure Intervention 4240
Weinan Wang, Li Zhu, Mehrab Bin Morshed, Md Mahbubur Rahman, Jungmok Bae, Jilong Kuang

Dynamic Language Group-Based MoE: Enhancing Code-Switching Speech Recognition with Hierarchical Routing 4245
Hukai Huang, Shenghui Lu, Yahui Shan, He Qu, Fengrun Zhang, Wenhao Guan, Qingyang Hong, Lin Li

Domain-Aware Node Representation Learning for Graph Out-Of-Distribution Generalization 4250
Yi Qiao, Yang Liu, Qing He, Xiang Ao

Quantum-Behaved Particle Swarm Optimization for the Segmentation of Kidney Stone CT Images 4255
Sajad Ahmad Rather, Akhilesh Kandwal, Mohammad Khalid Pandit, Partha Pratim Roy

Spectrum Blind Unlimited Sampling of Multi-Band Signals 4260
Ruiming Guo, Ayush Bhandari

On Noise-Sensitivity of Unlimited Sampling in Line Spectral Estimation 4265
Wenlong Wang, Zai Yang

CGDD: Contrastive Gaussian-Dirac Diffusion Model 4270
Chih-Chun Chen, Hsin-Yi Lin, Jen-Tzung Chien

DCCT-Net: A Network Combined Dynamic CNN and Transformer for Image Compressive Sensing 4275
Lijuan Xu, Haixiao Mei, Fenghua Tong, Dawei Zhao, Fuqiang Yu

Foreground-Aware Prototypical Network for Prohibited Item Detection from X-Ray Scans 4280
Xu Yang, Yufei Li, Long Tian, Haonan Shi, Ting Lan, Xiyang Liu

Decoupling While Coupling: Towards More Accurate Stereo Image Sand Removal Beyond Certainty 4285
Bingcai Wei, Hui Liu, Chuang Qian, Yi Jia, Wangyu Wu, Zhishan Li

Advances in Microphone Array Processing and Multichannel Speech Enhancement 4290
Gongping Huang, Jesper R. Jensen, Jingdong Chen, Jacob Benesty, Mads G. Christensen, Akihiko Sugiyama, Gary Elko, Tomas Gaensler

Multi-Source Data Lossless Compression Via Parallel Expansion Mapping and xLSTM 4295
Huidong Ma, Hui Sun, Liping Yi, Xiaoguang Liu, Gang Wang

Semi-VFL: Communication-Efficient Few-Label Vertical Federated Learning with Stacked Generalization and Model-Level Consistency 4300
Xuan Jin, Yuanzhi Yao, Caihong Kai, Rui Wang, Nenghai Yu

Class-Difficulty Aware Hybrid Active Learning 4305
Yaling Ge, Xun Pu, Jun Zhou

Multi-Scale Context Intertwining for Panoramic Renal Pathology Segmentation 4310
Ye Zhang, Xianchao Guan, Hengrui Li, Xiangming Yan, Ziyue Wang, Yongbing Zhang

A Fuzzy C-Means Clustering Algorithm for Real Medical Image Segmentation 4315
Feifei Zhang, Fei Shi, Dayong Ren, Yue Li

Learning Class Prototypes for Visual Emotion Recognition 4320
Jiankun Zhu, Sicheng Zhao, Jing Jiang, Zhaopan Xu, Wenbo Tang, Hongxun Yao

Naturalistic Music Decoding from EEG Data Via Latent Diffusion Models 4325
Emilian Postolache, Natalia Polouliakh, Hiroaki Kitano, Akima Connelly, Emanuele Rodolà, Luca Cosmo, Taketo Akama

Slungt: Even Faster Spoken Language Understanding with N-Grams and Tries 4330
Daniel Bermuth, Wolfgang Reif

Monte Carlo Score Matching for Image Generation 4335
Nishanth Shetty, Chandra Sekhar Seelamantula

Hybrid Pseudo-Labeling for Semi-Supervised Automatic Speech Recognition 4340
Lin Zheng, Han Zhu, Chengxu Yang, Xuyang Wang, Gaofeng Cheng, Ta Li

IReWindColor: Vision Transformer with Residual Embedding and Window Encoder for Point-Interactive Image Colorization 4345
Hideyuki Ogura, Masaaki Ikehara

A Comparative Analysis of Generalised Echo and Interference Cancelling and Extended Multichannel Wiener Filtering for Combined Noise Reduction and Acoustic Echo Cancellation 4350
Arnout Roebben, Toon van Waterschoot, Marc Moonen

Accelerating Computation for Large-Scale Wide-Band RF Imaging 4355
Ziyu Zhou, Yiming Zhou, Wei Dai

TFS: Revisiting Temporal Language Grounding from Frequency Spiking Perspective 4360
Yifan Lyu, Hongzhou Wu, Lixiang Liu, Chuxiong Sun

Maintaining Prosodic Consistency in Automatic Dubbing for Better Isochrony 4365
Parnia Bahar, Alejandro Pérez, Javier Iranzo-Sánchez

3D TDOA-AOA Quaternion Based Acoustic SLAM for Drone Localization and Source Mapping 4370
Hala Abualsaud, Peter Gerstoft

Low Frame-Rate Speech Codec: A Codec Designed for Fast High-Quality Speech LLM Training and Inference 4375
Edresson Casanova, Ryan Langman, Paarth Neekhara, Shehzeen Hussain, Jason Li, Subhankar Ghosh, Ante Jukić, Sang-Gil Lee

Geodesic Mean Threshold Scheme on Riemannian Manifold for EEG Signal Classification 4380
Srikireddy Dhanunjay Reddy, Tharun Kumar Reddy

COCO-OLAC: A Benchmark for Occluded Panoptic Segmentation and Image Understanding 4385
Wenbo Wei, Jun Wang, Abhir Bhalerao

Planing it by Ear: Convolutional Neural Networks for Acoustic Anomaly Detection in Industrial Wood Planers 4390
Anthony Deschênes, Rémi Georges, Cem Subakan, Bruna Ugulino, Antoine Henry, Michael Morin

Deep Unfolding Using Score-Based Generative Networks for Automotive Radar Interference Mitigation 4395
Xinyi Wei, Jihwan Youn, Jeroen Overdevest, Jun Li, Satish Ravindran, Ruud J.G. van Sloun

Perceptual Audio Coding: A 40-Year Historical Perspective 4400
Jürgen Herre, Schuyler Quackenbush, Minje Kim, Jan Skoglund

Enhancing Low-Resource ASR Through Versatile TTS: Bridging the Data Gap 4405
Guanrou Yang, Fan Yu, Ziyang Ma, Zhihao Du, Zhifu Gao, Shiliang Zhang, Xie Chen

Dual-Path Consistency Unsupervised Domain Adaptation for Nighttime Semantic Segmentation 4410
Yuwu Lu, Jicong Lang, Meirong Ding

DUNE: Sim2Real Transfer for Depth-Based Navigation in Unstructured Dynamic Indoor Environments 4415
Chaoyi Xu, Wen Liu, Jilong Wang, Lei Ma, Fan Yin, Zhongliang Deng

Multi-Level Speaker Representation for Target Speaker Extraction 4420
Ke Zhang, Junjie Li, Shuai Wang, Yangjie Wei, Yi Wang, Yannan Wang, Haizhou Li

MST-HA: Multi-Modal Signal Fusion with Bayesian Optimization for Robust Industrial Robot Joint Health Assessment 4425
Haoyu Wang, Zilong Yin, Bin Chen, Xiyue Yan, Chenyu Zhou, Beibei Zhang, Xinyuan Li, Guangmeng Xue, Haichao Xu

A Hierarchical Taxonomy for Deep State Space Models 4430
Shiqin Tang, Pengxing Feng, Shujian Yu, Yining Dong, S. Joe Qin

Subtractive Training for Music Stem Insertion Using Latent Diffusion Models 4435
Ivan Villa-Renteria, Mason Long Wang, Zachary Shah, Zhe Li, Soohyun Kim, Neelesh Ramachandran, Mert Pilanci

Impairments Are Clustered in Latents of Deep Neural Network-Based Speech Quality Models 4440
Fredrik Cumlin, Xinyu Liang, Victor Ungureanu, Chandan K. A. Reddy, Christian Schüldt, Saikat Chatterjee

Wave-U-Mamba: An End-To-End Framework for High-Quality and Efficient Speech Super Resolution 4445
Yongjoon Lee, Chanwoo Kim

DiffKillR: Killing and Recreating Diffeomorphisms for Cell Annotation in Dense Microscopy Images 4450
Chen Liu, Danqi Liao, Alejandro Parada-Mayorga, Alejandro Ribeiro, Marcello DiStasio, Smita Krishnaswamy

VITRO: Vocabulary Inversion for Time-Series Representation Optimization 4455
Filippos Bellos, Nam H. Nguyen, Jason J. Corso

Enhancing TTS Stability in Hebrew Using Discrete Semantic Units 4460
Ella Zeldes, Or Tal, Yossi Adi

Exploring Antenna Placement Configurations with a Radar-Based Silent Speech Interface 4465
João Menezes, Martin Schütze, Petr Schaffer, Dirk Plettemeier, Peter Birkholz

Know Your Heart Better: Multimodal Cardiac Output Monitoring Using Earbuds 4470
Hao Zhou, Md Mahbubur Rahman, Mehrab Bin Morshed, Yunzhi Li, Md Saiful Islam, Larry Zhang, Jungmok Bae, Christina Rosa, Wendy Berry Mendes, Jilong Kuang

Efficient Supernet Training with Orthogonal Softmax for Scalable ASR Model Compression 4475
Jingjing Xu, Eugen Beck, Zijian Yang, Ralf Schlüter

A Distillation-Based Future-Aware Graph Neural Network for Stock Trend Prediction 4480
Zhipeng Liu, Peibo Duan, Mingyang Geng, Bin Zhang

Contextual ASR with Retrieval Augmented Large Language Model 4485
Cihan Xiao, Zejiang Hou, Daniel Garcia-Romero, Kyu J Han

LLM-Powered Grapheme-To-Phoneme Conversion: Benchmark and Case Study 4490
Mahta Fetrat Qharabagh, Zahra Dehghanian, Hamid R. Rabiee

ANASETC: Automatic Neural Architecture Search for Encrypted Traffic Classification 4495
Heng Zhang, Ziqian Chen, Wei Xia, Gang Xiong, Gaopeng Gou, Zhen Li, Guangyan Huang, Yunpeng Li

TriFormer: Triple Branch Heterogeneous Transformers for UHD Image Restoration 4500
Xinchen Ma, Yue Lu

Optimization of Chirality Variation in Carbon Nanotube Field Effect Transistor Spiking Neurons 4505
Shelby Williams, Prosen Kirtonia, Kasem Khalil, Magdy Bayoumi

GCS-M3VLT: Guided Context Self-Attention Based Multi-Modal Medical Vision Language Transformer for Retinal Image Captioning 4509
Teja Krishna Cherukuri, Nagur Shareef Shaik, Jyostna Devi Bodapati, Dong Hye Ye

DiffETM: Diffusion Process Enhanced Embedded Topic Model 4514
Wei Shao, Mingyang Liu, Linqi Song

Dense-Sparse Dynamic Time Warping for Customizing Piano Concerto Accompaniments 4519
Tj Tsai, Kavi Dey, Yigitcan Özer, Meinard Müller

Past, Present, and Future of Spatial Audio and Room Acoustics 4524
Shoichi Koyama, Enzo De Sena, Prasanga Samarasinghe, Mark R. P. Thomas, Fabio Antonacci

Hypothesis Clustering and Merging: Novel MultiTalker Speech Recognition with Speaker Tokens 4529
Yosuke Kashiwagi, Hayato Futami, Emiru Tsunoo, Siddhant Arora, Shinji Watanabe

Online Proximal ADMM for Graph Learning from Streaming Smooth Signals 4534
Hector Chahuara, Gonzalo Mateos

Evidential-TTS: High Fidelity Zero-Shot Text-To-Speech Using Evidential Deep Learning 4539
Myeonghun Jeong, Minchan Kim, Semin Kim, Nam Soo Kim

Reliable Imputed-Sample Assisted Vertical Federated Learning 4544
Yaopei Zeng, Lei Liu, Shaoguo Liu, Hongjian Dou, Baoyuan Wu, Li Liu

Label Dependency Aware Loss for Reliable Multi-Label Medical Image Classification 4549
Aditya Shankar Pal, Arkapal Panda, Utpal Garain

Exploring Group Theory for Optimal Cognitive Radar Waveform Design 4554
Jonathan Monsalve, Kumar Vijay Mishra, A. Robert Calderbank

Mel-Spectrogram Inversion Via Alternating Direction Method of Multipliers 4559
Yoshiki Masuyama, Natsuki Ueno, Nobutaka Ono

DirichNet Model for Detection of TMS-Induced Speech Errors in Patients Undergoing Epilepsy Surgery 4564
Kodali Radha, Shalini Narayana

RoRA: Efficient Fine-Tuning of LLM with Reliability Optimization for Rank Adaptation 4569
Jun Liu, Zhenglun Kong, Peiyan Dong, Xuan Shen, Pu Zhao, Hao Tang, Geng Yuan, Wei Niu, Wenbin Zhang, Xue Lin, Dong Huang, Yanzhi Wang

Rejection of Workers with Heterogeneous (Mismatched) Data in Federated Learning * 4574
Rick S. Blum, Brian M. Sadler

Epigraph Based Multilevel Optimization (EMO) for Enhancing Chain-Of-Thought Reasoning Capabilities 4579
Songtao Lu, Yanna Ding, Lior Horesh, Jianxi Gao, Malik Magdon-Ismail

What Are They Doing? Joint Audio-Speech Co-Reasoning 4584
Yingzhi Wang, Pooneh Mousavi, Artem Ploujnikov, Mirco Ravanelli

Network Games Induced Prior for Graph Topology Learning 4589
Chenyue Zhang, Shangyuan Liu, Hoi-To Wai, Anthony Man-Cho So

Supervised Dimension Reduction Through Linear Projection 4594
Biao Chen, Joshua Kortje

Distributed ATC Particle Filters for Cooperative Quaternion Tracking 4599
Claudio J. Bordin, Marcelo G. S. Bruno, Stiven S. Dias

MAP Image Recovery with Guarantees Using Locally Convex Multi-Scale Energy (LC-MUSE) Model 4604
Jyothi Rikhab Chand, Mathews Jacob

Near Optimal Privacy Preserving Fair Multi-Agent Bandits 4609
Amir Leshem

FeedbackFuzz: Fuzzing Processors Via Intricate Program Generation with Feedback Engine 4614
Jiashun Wang, Baojiang Cui, Renhai Dong, Rundi Zhai

ALIC: Adaptive Fusion Entropy Model for Learned Image Compression 4619
Lingxue Li, Meiqin Liu, Yifan Zhang, Qi Tang, Chao Yao, Yao Zhao

Larger Language Models Don't Care How You Think: Why Chain-Of-Thought Prompting Fails in Subjective Tasks 4624
Georgios Chochlakis, Niyantha Maruthu Pandiyan, Kristina Lerman, Shrikanth Narayanan

MmWave-Whisper: Phone Call Eavesdropping and Transcription Using Millimeter-Wave Radar 4629
Suryoday Basak, Abhijeeth Padarthi, Mahanth Gowda

Unsupervised Image-To-Image Style Transfer Via Dual-Condition Diffusion Models* 4634
Anfei Fan, Jun Yang, Wei Li, Chiyu Zhang

SGRAND: Stochastic Graph Neural Diffusion 4639
Kaihang Dou, Fan Li, Suixiang Gao

Online Contrastive Continual Learning with Hard Negative Samples 4644
Guanglu Wang, Xinyue Liu, Wentao Yang, Wenxin Liang, Linlin Zong, Xianchao Zhang

Bridging Modality Gap with Large Speech and Language Models for End-To-End Speech-To-Text Translation 4649
Weitai Zhang, Simran Naagar, Zhongyi Ye, Peiwang Tang, Xinyuan Zhou, Junhua Liu, Lirong Dai

A Multi-Agent Multi-Environment Mixed Q-Learning for Partially Decentralized Wireless Network Optimization 4654
Talha Bozkus, Urbashi Mitra

DATA-VSR: Dynamic Trajectory Attention and Texture Adaptive Rooter for Video Super-Resolution 4659
Linfeng He, Meiqin Liu, Qi Tang, Chao Yao, Yao Zhao

Robust Dwell Time Allocation for Multiple Ballistic Reentry Target Tracking in Phased Array Radar 4664
Zishi Zhang, Ye Yuan, Wei Yi, Pramod K. Varshney

Adapt and Feature Translation for Class-Incremental Learning with Pre-Trained Models 4669
Hongfeng Li, Geming Xia, Yuze Zhang, Hongcheng Li, Hongwei Huang, Jiawen Wu

A Transmitter-Model Unaware Generative Image Compression Framework for Semantic Communication 4674
Rongcan Zheng, Xiaodan Song, Xuguang Zuo, Minxi Yang, Dahua Gao, Xuemei Xie

Causal Feature Supervision Decoupling: A Novel Method for Clothes-Changing Person Re-Identification Algorithm 4679
Wenxin Hu, Caidan Zhao, Chenxing Gao, Zhiqiang Wu

Maximum Likelihood Estimation for Bivariate Joint Distribution Recovery from Max-Aggregated Data 4684
Tianjian Zhang, Feng Yin, Yue Sun, Qi Yan

TSP: Task-Specific Pruning for Personalized Image Classification on Edge Devices 4689
Yanting Wang, Bojie Shi, Han Zhang

DBCR: Exploiting Both Intra-Cluster and Extra-Cluster Relations for Compositional Reasoning 4694
Chengtai Li, Guosheng Su, Jianfeng Ren, Ruibin Bai, Yitian Zhao, Xudong Jiang

RF-Pose Estimation Based on Contrastive Camera-Radar-Images Pretraining 4699
Yen-Hsiang Tseng, Po-Hsuan Tseng

Test-Time Alignment-Enhanced Adapter for Vision-Language Models 4704
Baoshun Tong, Kaiyu Song, Hanjiang Lai

Latent Diffusion Bridges for Unsupervised Musical Audio Timbre Transfer 4709
Michele Mancusi, Yurii Halychanskyi, Kin Wai Cheuk, Eloi Moliner, Chieh-Hsin Lai, Stefan Uhlich, Junghyun Koo, Marco A. Martínez-Ramírez, Wei-Hsiang Liao, Giorgio Fabbro, Yuki Mitsufuji

Membership Encoding for Black-Box Neural Network Watermarking 4714
Hangwei Zhang, Fang-Qi Li, Shi-Lin Wang

Estimation of Doppler, Range, and Direction of Targets in Wideband Bistatic Automotive Radar 4719
Ali Moussa, Wei Liu

Enhancing Few-Shot Out-Of-Distribution Detection with Gradient Aligned Context Optimization 4724
Baoshun Tong, Kaiyu Song, Hanjiang Lai

Edge First: Edge-Guided Geometry for Superior 3D Roof Wireframe Reconstruction* 4729
Qiaoqiao Hao, Ting Han, Yujun Liu, Shangfeng Huang, Duxin Zhu, Jinhe Su, Yundong Wu, Guorong Cai

EPI-Mamba: State Space Model for Semantic Segmentation from Light Fields 4734
Yan Li, Jianan Chen, Qiong Wang, Jinshan Xu

Collaborative Association Network for Multi-View Multi-Human Association and Tracking Using Constraint Optimization and Object Search 4739
Zixuan Zhang, Fan Shi, Chen Jia, Meng Zhao, Xu Cheng

Spatio-Temporal Multi-Subgraph GCN for 3D Human Motion Prediction 4744
Jiexin Wang, Yiju Guo, Bing Su

Take Attention Inside: Neighbor Pair Graph Contrastive Learning 4749
Bisheng Tang, Xiaojun Chen, Shaopu Wang, Yuexin Xuan, Zhendong Zhao

PHMamba: Preheating State Space Models with Context-Augmented Features for Medical Image Segmentation 4754
Nuo Chen, Shaoyu Wang, Ran Lu, Wenxuan Li, Xiujin Shi

A Quality-Aware Sampling Framework for Efficient 3D Point Cloud Transmission 4759
Puyue Hou, Qi Yang, Yue Li, Yujie Zhang, Jianchao Yang, Yiling Xu, Tiejun Huang

Collaborative Inference Acceleration with Non-Penetrative Tensor Partitioning 4764
Zhibang Liu, Chaonong Xu, Zhenjie Lv, Zhizhuo Liu, Suyu Zhao

Boosting Stereo Image Noise Removal by Learning Uncertainty and Enriched Features 4769
Bingcai Wei, Hui Liu, Chuang Qian, Xinyu Ren, Xintong Xu

EM-MIAs: Enhancing Membership Inference Attacks in Large Language Models Through Ensemble Modeling 4774
Zichen Song, Sitan Huang, Zhongfeng Kang

ViolinDiff: Enhancing Expressive Violin Synthesis with Pitch Bend Conditioning 4779
Daewoong Kim, Hao-Wen Dong, Dasaem Jeong

Serial Local Patterns and Irregular Dependencies Extract and Cascaded Fusion Network for Structural Crack Segmentation 4784
Hui Liu, Chen Jia, Xu Cheng, Xiufeng Liu, Fan Shi

Temporal Dynamics Decoupling with Inverse Processing for Enhancing Human Motion Prediction 4789
Jiexin Wang, Yiju Guo, Bing Su

Norm Augmented Graph AutoEncoders for Link Prediction 4794
Yunhui Liu, Huaisong Zhang, Xinyi Gao, Liuye Guo, Zhen Tao, Tieke He

Keep What You Need : Extracting Efficient Subnetworks from Large Audio Representation Models 4799
David Genova, Philippe Esling, Tom Hurlin

Time-Graph Frequency Representation with Singular Value Decomposition for Neural Speech Enhancement 4804
Tingting Wang, Tianrui Wang, Meng Ge, Qiquan Zhang, Zirui Ge, Zhen Yang

FairAdapter: Detecting AI-Generated Images with Improved Fairness 4809
Feng Ding, Jun Zhang, Xinan He, Jianfeng Xu

Dynamic Graph Recommendation Via Sparse Augmentation and Singular Adaptation 4814
Zhen Tao, Yuehang Cao, Yang Fang, Yunhui Liu, Xiang Zhao, Tieke He

Robust Adversarial Training for Industrial Defect Classification with Long-Tailed Data 4819
Shuchun Xu, Jiguang Lyu, Dapeng Man, Hengheng Xiong, Tao Liu, Wu Yang

Unified Graph and Hypergraph Neural Network for Next-Item Recommendation 4824
Xinyu Zhang, Qize Jiang, Hanyuan Zhang, Weiwei Sun

Using RLHF to Align Speech Enhancement Approaches to Mean-Opinion Quality Scores 4829
Anurag Kumar, Andrew Perrault, Donald S. Williamson

A Domain Adaptation Framework for Speech Recognition Systems with Only Synthetic Data 4834
Minh Tran, Yutong Pang, Debjyoti Paul, Laxmi Pandey, Kevin Jiang, Jinxi Guo, Ke Li, Shun Zhang, Xuedong Zhang, Xin Lei

HDMoLE: Mixture of LoRA Experts with Hierarchical Routing and Dynamic Thresholds for Fine-Tuning LLM-Based ASR Models 4839
Bingshen Mu, Kun Wei, Qijie Shao, Yong Xu, Lei Xie

MF-BERT: A Siamese Pre-Training Framework for Motion Forecasting 4844
Jianxin Shi, Jinhao Chen, Xiaolong Chen, Jun Ma, Tianyu Wo

Efficient-USR: Prompt Guided Dual-Domain Feature Information for Efficient Underwater Image Super-Resolution 4849
Alik Pramanick, Utsav Bheda, Arijit Sur

FlowMAC: Conditional Flow Matching for Audio Coding at Low Bit Rates 4854
Nicola Pia, Martin Strauss, Markus Multrus, Bernd Edler

Mixed Spiking NeRF: Towards a More Efficient Neural Radiance Fields 4859
Kaian Wang, Longhao Zou

BLR-MoE: Boosted Language-Routing Mixture of Experts for Domain-Robust Multilingual E2E ASR 4864
Guodong Ma, Wenxuan Wang, Lifeng Zhou, Yuting Yang, Yuke Li, Binbin Du

Deep Transfer Regression for EEG-Based Driving Fatigue Detection 4869
Yikai Zhang, Yong Peng, Ziyue Yang, Feiwei Qin, Wanzeng Kong

SHAP-Integrated Convolutional Diagnostic Networks for Feature-Selective Medical Analysis 4874
Yan Hu, Ahmad Chaddad

Multi-Relational Geometric Regularization Framework for Multi-Modal Emotion Recognition in Conversation 4879
Tao Zhang, Zhenhua Tan

SCDiar: A Streaming Diarization System Based on Speaker Change Detection and Speech Recognition 4884
Naijun Zheng, Xucheng Wan, Kai Liu, Zhou Huan

Meta-Conscious Driven Domain-Aware Federated Learning 4889
Zilong Yin, Haoyu Wang, Xiaogang Lin, Xin Zhang, Bin Chen, Chenyu Zhou

Learning Semantic Facial Descriptors for Accurate Face Animation 4894
Lei Zhu, Yuanqi Chen, Xiaohang Liu, Thomas H. Li, Ge Li

Symmetric Bi-Branch Modality-Search Aggregation Network for Multi-Modal Liver Segmentation 4899
Huaxiang Liu, Jie Yang, Jie Jin, Youyao Fu, Shiqing Zhang, Wenbin Ji, Dandan Wang, Jiangxiong Fang

OLN++: Improved Object Localization Network for Open-World Object Detection 4904
Haonan Mai, Libo Huang, Zhulin An, Jiarui Zhao, Chuanguang Yang, Erhu Zhao, Yongjun Xu

Cooperative Multi-Target Tracking Based on Multi-Detection TPHD in MIMO-OFDM Systems 4909
Chen Zhong, Lan Tang, Ying-Chang Liang

One Shot is Enough for Sequential Infrared Small Target Segmentation 4914
Bingbing Dan, Meihui Li, Tao Tang, Jing Zhang

Exploring Acted Sleepy Speech to Advance Real-World Sleepiness Estimation and Cognitive Degradation Detection 4919
Jihye Moon, Youngsun Kong, Yashvi Gupta, Ki H. Chon

Influence-Based Channel Reweighting for Multivariate Time Series Forecasting 4924
Yipu Liu, Zheng Wang, Qinghua Hu

DSSM: Dual State Space Model for Human Motions Generation 4929
Yiming Liu, Huan Zhao, Yaqian Liu, Haijiao Chen, Bo Li, Guanghui Ye, Zixing Zhang

TGCA: A Transformer GNN-Based Approach with Cross-Attention Mechanism for Steganographic Text Detection in Social Networks 4934
Junkai Lu, Zhongliang Yang, Kaibo Huang, Zhuang Wang, Zhili Zhou, Linna Zhou

MncCap: Mining Neural Composition for Zero-Shot Image Captioning Via Text-Only Training 4939
Tongtong Liu, Chen Yang, Guoqiang Chen, Qinxu Gao, Enhua Song, Wenhui Li

Enhancing the Robustness of LiDAR-Based Object Detection Under Disappearing Attacks 4944
Huiying Wang, Lisong Zhang, Wenbo Wang, Yu Wen

Boosting Code-Switching ASR with Mixture of Experts Enhanced Speech-Conditioned LLM 4949
Fengrun Zhang, Wang Geng, Hukai Huang, Yahui Shan, Cheng Yi, He Qu

LKConvPose: A Pose Estimation Model with Large Receptive Field 4954
Ying Huang, Qiang Chen, Xiu-Xiu Zhan, Jianzhang Zhang, Chuang Liu

Lungmix: A Mixup-Based Strategy for Generalization in Respiratory Sound Classification 4959
Shijia Ge, Weixiang Zhang, Shuzhao Xie, Baixu Yan, Zhi Wang

Catch Causal Signals from Edges for Label Imbalance in Graph Classification 4964
Fengrui Zhang, Yujia Yin, Hongzong Li, Yifan Chen, Tianyi Qu

Confusion-Aware Prototypical Contrastive Learning for Open-Vocabulary Object Detection 4969
Hongqiang Cheng, Le Jiang, Guoming Li, Xiaozhou Ye, Ye Ouyang

Physically Robust and Imperceptible Adversarial Examples Generation Based on Frequency 4974
Chengyao Hua, Yitian Chen, Shigeng Zhang, Xuan Liu, Senzhang Wang, Weiping Wang, Kai Chen

LossControl: Defending Membership Inference Attacks by Controlling the Loss 4979
Bo Yang, Hongwei Yang, Renhao Lu, Hui He, Weizhe Zhang, Haoyu He, Rahul Yadav

SCAT: Shared-Convolution Adaptation Tuning for Foreground Segmentation 4984
Kaiwen Li, Dezheng Gao, Zelin Yang, Xing Wei

DepMamba: Progressive Fusion Mamba for Multimodal Depression Detection 4989
Jiaxin Ye, Junping Zhang, Hongming Shan

Efficient Localized Perception for Resource-Constrained Vision Systems 4994
A. V. Subramanyam, Niyati Singal, Vinay K Verma

A Bilinear Source Separation, Dereverberation, and Background Noise Suppression Algorithm for Augmented Reality Applications 4999
Alon Nemirovsky, Gal Itzhak, Israel Cohen

Efficient Visual Storytelling Through Descriptive Words Distillation and Dynamic Decoding 5004
Zixuan Xiong, Hai Lin, Lichen Bai, Yinghui Li, Hai-Tao Zheng, Hong-Gee Kim

HiE-VL: A Large Vision-Language Model with Hierarchical Adapter for Handwritten Mathematical Expression Recognition 5009
Hong-Yu Guo, Fei Yin, Jian Xu, Cheng-Lin Liu

LNeRV: Learnable Hierarchical Encoding Improve Neural Representation Video Codec 5014
Jiahong Chen, Xiang Liu, Bin Chen, Baoyi An, Dai Tao, Shu-Tao Xia

Quantum Neural Networks: A Path to Lower Emissions Through Fuel Consumption Prediction in Shipping 5019
S.F. Chien, Julien J.M. Hermans, Austin A. Kana, Charilaos. C. Zarakovitis, Stathis Zavvos, H.S. Lim

Parameter-Efficient Federal-Tuning Enhances Privacy Preserving for Speech Emotion Recognition 5024
Haijiao Chen, Huan Zhao, Yingxue Gao, Yiming Liu, Zixing Zhang

MFT: Modal Fusion Transformer for Cross-Modal Fusion in 3D Object Detection 5029
Haojie Cai, Dongfu Yin, Fei Yu, Siting Xiong

CJST: CTC Compressor Based Joint Speech and Text Training for Decoder-Only ASR 5034
Wei Zhou, Junteng Jia, Leda Sari, Jay Mahadeokar, Ozlem Kalinli

Multi-Degradation Oriented Deep Unfolding Model for Hyperspectral Image Reconstruction 5039
Xian-Hua Han, Jian Wang

Harnessing Content and Structure in ID for Multimodal Recommendation 5044
Yuting Liu, Enneng Yang, Yizhou Dang, Guibing Guo, Qiang Liu, Yuliang Liang, Linying Jiang, Xingwei Wang

EfficientNet-Gaze: Integrating Multi-Scale Feature Extraction with Frequency Domain Analysis for Efficient Gaze Estimation 5049
Yanxia Wang, Guoyu Xia

AMuSE: Attentive Multilingual Speech Encoding for Zero-Prior ASR 5054
Ashutosh Varshney, Debmalya Chakrabarty, Akshat Jaiswal, Harish Arsikere, Abhinav Jain, Swayambhu Nath Ray, Frederick Weber, Anand Mohan, Prantik Sen, Garima Lalwani, Sambuddha Bhattacharya, Sri Garimella

Simultaneous Music Separation and Generation Using Multi-Track Latent Diffusion Models 5059
Tornike Karchkhadze, Mohammad Rasool Izadi, Shlomo Dubnov

Robust Hybrid Convolutional Beamspace for High Resolution DOA Estimation 5064
Rubin Jose Peter, Sooraj K. Ambat

Non-Invasive Speaker-Dependent Continuous Phoneme Recognition with a Radar-Based Silent Speech Interface 5069
João Menezes, Christoph Wagner, Peter Steiner, Petr Schaffer, Dirk Plettemeier, Peter Birkholz

VCSA: Video Copy Localization Via Single Frame Annotation 5074
Peng Li, Shuguo Hu, Yang Yang, Huaiwen Zhang

Spatial Frequency-Aware Self-Distillation for Weakly-Supervised Semantic Segmentation 5079
Jingyuan Fang, Yang Ning, Xiushan Nie

Delving into Coarse-Fine Feature Interaction Alignment for UAV Object Detection 5084
Yanchao Bi, Yang Ning, Xiushan Nie

Inference Retrieval-Augmented Multi-Modal Chain-Of-Thoughts Reasoning for Language Models 5089
Qiangqiang He, Shuwei Qian, Jie Zhang, Chongjun Wang

Bone Conducted Signal Guided Speech Enhancement for Voice Assistant on Earbuds 5094
Jens Heitkaemper, Joe Caroselli, Max McKinnon, Arun Narayanan, Nathan Howard

Learning-Aided Kalman Tracking in Biased Dynamic Systems: The Case of Cable-Driven Robots for Surgery 5099
Linoy Ketashvili, Shachar Ashkenasy, Ilana Nisky, Nir Shlezinger

Learning Rate Optimization for Deep Neural Networks Using Lipschitz Bandits 5104
Padma Priyanka, Sheetal Kalyani, Avhishek Chatterjee

RFEM: Remote Feature Enhancement Module for Target Detection 5109
Feng Hu, Chuangye Wang, Jian Xiong, Wenhua Zhang, Haolun Li, Hao Gao

ControlMol: Adding Substructure Control to Molecule Diffusion Models 5114
Zhengyang Qi, Zijing Liu, Jiying Zhang, He Cao, Xiaohua Xu, Yu Li

Design and Optimization of Superdirective Beamforming and Post-Filtering for Speech Enhancement 5119
Xiaoran Yang, Gongping Huang, Jilu Jin, Jingdong Chen, Jacob Benesty

Improving Sidescan Sonar Performance Using Array Upsampling Beamforming Synthetic Aperture 5124
Weibo Mao, Dongdong Zhao, Peng Chen, Shihui Liang, Yiran Li, Ronghua Liang

EP-SAM: An Edge-Detection Prompt SAM Based Efficient Framework for Ultra-Low Light Video Segmentation 5129
Zhitao Wang, Jiangtao Wen, Yuxing Han

VLIMNet: A Visible Light and Infrared Image Matching Network Based on Segment Anything Model and SuperPoint 5134
Zhongyuan Chen, Zhan Zhang, Decheng Zuo, Ning Wang, Liufeng Fan, Zhiwei Liu

COREMIL: Contextual Position Encoding-Based Retrievable Multiple Instance Learning for Slide-Level Classification 5139
Bingchen Li, Qiming He, Junru Cheng, Tian Guan, Yonghong He, Guangde Zhou

Vision-Text Enhancement Network for Weakly Supervised Video Anomaly Detection 5144
Yiheng Chen, Shuai Fu, Niantong Qin, Xinning Du, Jianping Ren, Shuhua Liu

Visual Entity-Centric Prompting for Knowledge Retrieval in Knowledge-Based VQA 5149
Jiaming Yang, Jiuxiang You, Ziyue Qiu, Guobo Xie, Yi Yu, Zhenguo Yang

Learning Adaptive Spatial-Temporal Structured Correlation Filters for UAV Object Tracking 5154
Gong Li, Wei Zhao, Peng Liu, Xianglong Tang

ZSVC: Zero-Shot Style Voice Conversion with Disentangled Latent Diffusion Models and Adversarial Training 5159
Xinfa Zhu, Lei He, Yujia Xiao, Xi Wang, Xu Tan, Sheng Zhao, Lei Xie

FreeSegDiff: Annotation-Free Saliency Segmentation with Diffusion Models 5164
Chaofan Ma, Yuhuan Yang, Chen Ju, Yue Shi, Ya Zhang, Yanfeng Wang

EqGAN: Reformation-Based Feature Equalization Fusion for Few-Shot Image Generation 5169
Yingbo Zhou, Zhihao Yue, Yutong Ye, Pengyu Zhang, Xian Wei, Mingsong Chen

CE-FFT: Communication-Efficient Federated Fine-Tuning for Large Language Models Via Quantization and In-Context Learning 5174
Pengyu Zhang, Yingjie Liu, Yingbo Zhou, Ming Hu, Xian Wei, Mingsong Chen

Make Unseen Clear: Occluder Removal for Complete 3D Pedestrian Detection 5179
Yan Luo, Zefeng Qian, Chongyang Zhang, Yi Wang, Xiaolei Hu

Towards Detecting LLMs Hallucination Via Markov Chain-Based Multi-Agent Debate Framework 5184
Xiaoxi Sun, Jinpeng Li, Yan Zhong, Dongyan Zhao, Rui Yan

Global Enhanced Frame Prompt Tuning for Sound Event Detection 5189
Shiyu Yu, Lijian Gao, Qirong Mao

FiTGAN: Content Fusion with Style Transformation for Few-Shot Image Generation 5194
Yingbo Zhou, Pengyu Zhang, Yutong Ye, Zhihao Yue, Xian Wei, Mingsong Chen

MFANet: Multi-Feature Aggregation Network for Multi-Focus Image Fusion 5199
Libo Zhao, Xiaoli Zhang, Bo Huang, Mingjie Tian, Zeyu Wang

Adaptive Sparse Feature Location Activation Strategy for Sparse Detectors on Drone Images 5204
Yixuan Li, Yulong Xu, Pengnian Wu, Xuqi Yang, Meng Zhang

PAUSE: Privacy-Aware Active User Selection for Federated Learning 5209
Ori Peleg, Natalie Lang, Stefano Rini, Nir Shlezinger, Kobi Cohen

Ultra-Low Latency Speech Enhancement -A Comprehensive Study 5214
Haibin Wu, Sebastian Braun

Fine-Grained Global Modeling Learning for Personalized Federated Sequential Recommender 5219
Yicheng Di, Hongjian Shi, Ruhui Ma, Yuan Liu, Weiyu Wang

Symbol-Level Precoding-Based Self-Interference Cancellation for ISAC Systems 5224
Shu Cai, Zihao Chen, Ya-Feng Liu, Jun Zhang

DuPI: Dual-Resolution Pseudo-Label Integration for Semi-Supervised Instance Segmentation 5229
Yue Ma, Jie Hu, Chen Chen, Shengchuan Zhang, Xianming Lin, Liujuan Cao

Supervisor Alignment Framework: Enhancing LLM Alignment with Query-Ignoring Strategy and Multi-Agent Interaction 5234
Ziqun Bao, Yu Ji, Wen Wu, Xi Chen, Liang He

Face Relighting with Ratio Function for Explicit Geometric Representation 5239
Yiyang Hu, Zequn Zhang, Hui Zhang, Guquan Jing, Peng Gao

Transforming Classification with Federated Learning on Blockchain: A Unique Model Integration Approach 5244
Zhihao Hao, Bob Zhang, Haisheng Li

Reexamining the Efficacy of MetricGAN for Speech Enhancement 5249
Haibin Wu, Ali Aroudi, Buye Xu, Ashutosh Pandey, Francesco Nesta, Anurag Kumar, Alexander Reich, Ke Tan

Time-Domain Beamforming for Room Acoustics Analysis Based on Reverberant Field Estimation 5254
Tatiana Gelvez-Barrera, Quentin Leclere, Barbara Nicolas, Jérôme Antoni, Adrian Basarab

A Method for Removing Reflections from Water Surface Images Based on Pre-Trained Image Restoration 5259
Minghua Zhao, Rui Zhi, Shuangli Du, Jing Hu, Cheng Shi, Lin Wang

Exploiting Wavelet Scattering Transform & Squeeze-Excitation Blocks with Cross-Modal Attention for Multi-Modal Emotion Recognition 5264
Junchen Liu, Jesin James, Karan Nathwani

SSC: 106 Bit/s Ultra-Low Bitrate Semantic Speech Coding 5269
Renjie Jia, Zhiqiang He, Kai Niu, Zixuan Xiao, Yonghui Liu, Jianbing Liu

Easy-To-Hard Instance-Level Feature Fusion for Co-Saliency Detection 5274
Chuang Ding, Zhidong Han, Gang Dong, Lingyan Liang, Dongchao Wen, Kaihua Zhang

Learning Joint Appearance and Shape Co-Representations for Co-Saliency Detection 5279
Guanting Guo, Shenglong Hu, Huihui Song, Kaihua Zhang

Low-Rank Tucker Decomposition of Multi-Subject Complex-Valued fMRI Data 5284
Bin-Hua Zhao, Qiu-Hua Lin, Yue Han, Jia-Yang Song, Yan-Wei Niu, Vince D. Calhoun

LCE: A Framework for Explainability of Ultrasound Image Based on Concept Discovery 5289
Weiji Kong, Xun Gong, Juan Wang

Generative Adversarial Network with Adaptive Synthesis for Brain-Computer Interfaces in Motor Imagery Classification 5294
Qiaoli Zhou, Xiyuan Ye, Shurui Li, Yi Zhao, Qiang Du, Li Ke

Continuously Learning Video-Level Object Tokens for Robust UAV Tracking 5299
Bin Chen, Shenglong Hu, Gang Dong, Lingyan Liang, Dongchao Wen, Kaihua Zhang

Data-Driven White Noise Gain Constrained Robust Superdirective Beamformer for Speech Enhancement 5304
Hanchen Pei, Gongping Huang, Jilu Jin, Jianbo Ma, Zhizheng Wu, Jingdong Chen, Jacob Benesty

Large Language Models Are Strong Audio-Visual Speech Recognition Learners 5309
Umberto Cappellazzo, Minsu Kim, Honglie Chen, Pingchuan Ma, Stavros Petridis, Daniele Falavigna, Alessio Brutti, Maja Pantic

Feature Disentangling Dual-Stream Network for User Bias Alleviation in Social Media Prediction 5314
Wenhao Hu, Weilong Chen, Weimin Yuan, Xiaolu Chen, Han Yang, Yanru Zhang, Zhu Han

Open-Vocabulary Saliency-Guided Progressive Refinement Network for Unsupervised Video Object Segmentation 5319
Zhidong Han, Shenglong Hu, Huihui Song, Kaihua Zhang

Classification of Zhuang Dialect Combined with Bert and SimAM 5324
Min Huang, Xuejun Zhang, Wenkang Chen

Eye Movements as Images: A Multimodal Framework for Eye Movements Representation 5329
Dongsen Zhang, Peipei Li, Zekun Li, Yiwei Ru, Huijia Wu, Zhaofeng He

Generating Apoptosis-Inducing Anticancer Peptides Targeting BCL-XL Using Latent Diffusion Models on Small Datasets 5334
Tiara Natasha Binte Sayuti, Wang Conghao, Jagath C Rajapakse

Semi-Supervised Multilingual Alignment with Lexical Memory for Massively Parallel Text Mining 5339
Weitai Zhang, Peiwang Tang, Chao Lin, Simran Naagar, Zhongyi Ye, Junhua Liu

Harnessing Light Field Angular Cues and Spatial Geometries for Semantic Segmentation 5344
Chen Jia, Fan Shi, Xu Cheng

MetricGAN+KAN: Kolmogorov-Arnold Networks in Metric-Driven Speech Enhancement Systems 5349
Yemin Mai, Stefan Goetze

Spatio-Semantic Prompt Guided Adaptive Segment Anything for Remote Sensing Change Detection 5354
Shenglong Hu, Zhidong Han, Gang Dong, Lingyan Liang, Dongchao Wen, Kaihua Zhang

MKD-YOLO: Multi-Scale and Knowledge-Distilling YOLO for Efficient PPE Compliance Detection 5359
Juntao Zan, Yang Fang, Qilie Liu, Uswah Khairuddin, Yan Li, Kaiwei Sun

VPCI: Self-Supervised Visual Prompt-Guided Cross-Domain Interactive Image Fusion Framework 5364
Yong Liu, Chengyu Wu, Jiayuan Cui, Bin Jiang

ILDiff: Generate Transparent Animated Stickers by Implicit Layout Distillation 5369
Ting Zhang, Zhiqiang Yuan, Yeshuang Zhu, Jie Zhou, Jinchao Zhang

Learning Deep Frequency Degradation Prior for Remote Sensing Spatio-Temporal Fusion 5374
Yiting Bian, Shenglong Hu, Huihui Song, Kaihua Zhang

Enhancing Zero-Shot Relation Extraction Through Staged Interaction with Large Language Models 5379
Yifang Zhang, Pengfei Duan, Yiwen Yang, Shengwu Xiong

Enhancing Small Model Performance in Educational Classification Tasks Through Knowledge Distillation 5384
Haoxin Xu, Changyong Qi, Bingqian Jiang, Tong Liu, Longwei Zheng, Xiaoqing Gu

WaveSpect: A Hybrid Approach to Synthetic Audio Detection Via Waveform and Spectrogram Analysis 5389
Dong Chen, Fan Huang, Zhengxuan Song, Wei Zhu, Yin Yang, Kun Zeng

InstantSpeech: Instant Synchronous Text-To-Speech Synthesis for LLM-Driven Voice Chatbots 5394
Muyang Du, Chuan Liu, Junjie Lai

Large Language Models Are Efficient Learners as Zero-Shot Speech Translators 5399
Chenxuan Liu, Liping Chen, Peiwang Tang, Weitai Zhang, Xiaoxi Li, Sreyan Ghosh, Zhongyi Ye, Mingjia Yu

SUVAD: Semantic Understanding Based Video Anomaly Detection Using MLLM 5404
Shibo Gao, Peipei Yang, Linlin Huang

Group-Wise Semantic-Enhanced Interaction Network for Remote Sensing Spatio-Temporal Fusion..... 5409
Baoluo Zhu, Shenglong Hu, Huihui Song, Kaihua Zhang

DEBT: Enhancing Entity Alignment in Knowledge Graphs Through Description Enrichment and Bootstrap Training 5414
Ting Xiang, Jiapeng Zhang, Changjian Chen, Zhuo Tang

Deep Generic Representations for Domain-Generalized Anomalous Sound Detection 5419
Phurich Saengthong, Takahiro Shinozaki

Assessing Robustness of Multi-Modal Large Language Models in Image Classification Through Hierarchical WordNet-Based Evaluation..... 5424
Chang Liu, Hai Chen, Boxiang Wang, Shibao Zheng

CoF: Coarse to Fine-Grained Image Understanding for Multi-Modal Large Language Models..... 5429
Yeyuan Wang, Dehong Gao, Bin Li, Rujiao Long, Lei Yi, Xiaoyan Cai, Libin Yang, Jinxia Zhang, Shanqing Yu, Qi Xuan

Trainable Adaptive Score Normalization for Automatic Speaker Verification..... 5434
Jeong-Hwan Choi, Ju-Seok Seong, Ye-Rin Jeoung, Joon-Hyuk Chang

TD-GS: Few-Shot Object View Synthesis Via Task-Disentangled 3D Gaussian Splatting..... 5439
Jialin Wu, Ying Liu, Xu Wang, Xiaohao Zhang, Zhuo Tang, Ruihui Li

Generalize Audio Deepfake Algorithm Recognition Via Attribution Enhancement..... 5444
Zhigang Wang, Dengpan Ye, Jingyang Li, Jiacheng Deng

MMFN: Multi-Feature Multi-Modal Fusion Network for Diagnosis of Superficial Lymph Node Disease 5449
Yuankun Wang, Cheng Zhao, Yingxin Liu, Baiying Lei, Tianfu Wang, Luyao Zhou

VoxVietnam: A Large-Scale Multi-Genre Dataset for Vietnamese Speaker Recognition..... 5454
Hoang Long Vu, Phuong Tuan Dat, Pham Thao Nhi, Nguyen Song Hao, Nguyen Thi Thu Trang

Automatic Geometric Quantification and Rupture Risk Evaluation of 3D Intracranial Aneurysms 5459
Xudong Ru, Zeyao Zhang, Xingce Wang, Jing-Yi Liu, Yi-Cheng Zhu, Zhongke Wu

AudioCache: Accelerate Audio Generation with Training-Free Layer Caching 5464
Qingyang Shi, Zhicheng Du, Jiasheng Lu, Yingshan Liang, Xinyu Zhang, Yiran Wang, Jing Peng, Kehong Yuan

Planetary Gear Vibration Monitoring Using Synchronous Demodulation 5469
Rik Vaerenberg, Konstantinos Gryllias

UAV-Mounted SIM: A Hybrid Optical-Electronic Neural Network for DoA Estimation 5474
Shining Lin, Jiancheng An, Lu Gan, Mérouane Debbah

Noise Supervised Contrastive Learning and Feature-Perturbed for Anomalous Sound Detection..... 5479
Shun Huang, Zhihua Fang, Liang He

Multi-Scale Across Attention Incorporated Network for X-Ray Coronary Vessel Segmentation 5484
He Deng, Tong Fang, Xiangde Min

SuperPromptSeg: A Novel Fine-Tuning-Free Segmentation Method Leveraging Superpixel-Based Point Prompts 5489
Ziqiao Zhou, Min Cen, Yuzhe Zhang, Hong Zhang, Xu Steven Xu

TSCheater: Generating High-Quality Tibetan Adversarial Texts Via Visual Similarity 5494
Xi Cao, Quzong Gesang, Yuan Sun, Nuo Qun, Tashi Nyima

MSRFormer: Hybrid Scale Self-Attention and Local Fast Convolution Transformer for Facial Expression Recognition..... 5499
Zi-Qiang Shen, Yu-Yi Tang, Jun-Feng Yan, Yang Li, Guo-Ying Zhao

BlurPaint: Image Inpainting Using Blurring Diffusion Models 5504
Linxu Chen, Zhiqing Guo, Liejun Wang, Ke Lu

BIF: A Biosignature Identification Framework for Model-Agnostic Interpretation of MVI Diagnosis Models in HCC..... 5509
Pengyu Zheng, Yong Yi, Ye Luo

DARNet: A Dual Attention Residual Network for Medical Image Classification..... 5514
Ao Zhang, Zhenghua Guan, Tengda Zhang, Wenzheng Hu, Yi Liu, Baiying Lei

Near-Field ISAC in 6G: Addressing Phase Nonlinearity Via Lifted Super-Resolution..... 5519
Sajad Daei, Amirreza Zamani, Saikat Chatterjee, Mikael Skoglund, Gabor Fodor

SFE-Net: Harnessing Biological Principles of Differential Gene Expression for Improved Feature Selection in Deep Learning Networks..... 5524
Yuqi Li, Yuanzhong Zheng, Yaoxuan Wang, Jianjun Yin, Haojun Fei

MSE-Based Sampling of Bandlimited Product Graph Signals Via Joint Low-Pass Impulse Responses..... 5529
Fen Wang, Baoyi Xu, Xuyao Kang, Peng Ren, Long Yang

Stable Extended U-Net for Noise-Robust Speaker Verification..... 5534
Zonghui Wang, Zhihua Fang, Liang He

LINK: Adaptive Modality Interaction for Audio-Visual Video Parsing..... 5539
Langyu Wang, Bingke Zhu, Yingying Chen, Jinqiao Wang

So Far Yet So Near: Time Series Data Augmentation with Exploring non-Semantic Boundaries Based on Reinforcement Learning 5544
Haoran Li, Zhibo Zhang, Jiarong Kang, Xun Jiang, Xiaoli Gong, Jin Zhang, Zhe Sun, Andrzej Cichocki

Exploring Text-Queried Sound Event Detection with Audio Source Separation 5549
Han Yin, Jisheng Bai, Yang Xiao, Hui Wang, Siqi Zheng, Yafeng Chen, Rohan Kumar Das, Chong Deng, Jianfeng Chen

DiffMEL: A Large-Scale Difficulty-Graded Dataset for Multimodal Entity Linking 5554
Fang Wang, Xiaoying Bai, Tianwei Yan, Minghao Hu, Yi Liang

Emotional Knowledge Self-Distillation in Dialogue 5559
Zhongquan Jian, Yancheng Wang, Weichao Wu, Junfeng Yao, Meihong Wang, Qingqiang Wu

Communication-Efficient Exact Diffusion for Decentralized Learning 5564
Gustavo Faia, Stefan Vlaski, Roula Nassif

Essentia: Boosting Artifact Removal from EEG Through Semantic Guidance Utilizing Diffusion Model 5569
Haoran Li, Zhibo Zhang, Yuchen Li, Xiaoli Gong, Jin Zhang, Tingjuan Lu, Jin Zhou, Zhe Sun, Andrzej Cichocki

Attention-Based Beamformer for Multi-Channel Speech Enhancement 5574
Jinglin Bai, Hao Li, Xueliang Zhang, Fei Chen

ADC-GS: Pose-Free 3D Gaussian Splatting with Adaptive Depth Consistency 5579
Runling Liu, Shihe Shen, Na Jiang, Jiahao Wu, Guanhua Wu, Zhiyan Wang, Lu Xiao, Zhanke Wang, Ronggang Wang

KAFQN: Kolmogorov-Arnold Fuzzy-Guided Q-Network in Reinforcement Learning 5584
Bo Zhao, Zhizhong Liu, Zhuo Tang

ASANet:Scene Text Recognition with Alternate Self-Attention 5589
Wenting Xu, Elham Eli, Alimjan Aysa, Xuebin Xu, Kurban Ubul

GIST: Guided Interpretable Large Language Model Strategy Transfer for Multi-Task Reinforcement Learning 5594
Bo Zhao, Zhuo Tang

Investigation of Perceptual Music Similarity Focusing on Each Instrumental Part 5599
Yuka Hashizume, Tomoki Toda

BCS-Net: Multi-Task Breast Cancer Screening Network Enhanced by Multi-Modality Attention 5604
Ruili Li, Ruiyu Li, Eichi Takaya, Zizhen Lin, Tomoya Kobayashi, Nanako Mtsuda, Takuya Ueda

Lenna: Language Enhanced Reasoning Detection Assistant 5609
Fei Wei, Xinyu Zhang, Ailing Zhang, Bo Zhang, Xiangxiang Chu

Dynamic SRM Curriculum for Trustworthy Multi-Modal Classification 5614
Jian Zhu, Cui Yu, Xin Zou, Zhangmin Huang, Chenshu Hu, Jun Sun, Bo Lyu, Lei Liu, Chang Tang, Li-Rong Dai

TAGMO: Temporal Control Audio Generation for Multiple Visual Objects Without Training 5619
Xinyu Zhang, Keyu Fan, Yiran Wang, Yingshan Liang, Jiasheng Lu, Zhicheng Du, Qingyang Shi, Peiwu Qin

A Triangular Stable Node Network Based on Self-Supervised Learning for Personalized Prediction 5624
Qing Liu, Qian Gao, Jun Fan, Zhiqiang Zhang

Dual-Process Watermarked Diffusion: Integrating Watermarking with Denoising in Point Clouds 5629
Jinfu Wei, Heng Chang, Xiaohang Liu, Li Liu, Likun Li, Shiji Zhou, Chengyuan Li, Di Xu, Wei Gao, Ran Liao

Self-Supervised Speaker Verification with Batch-Scale Pseudo-Labels Correction 5634
Junxu Wang, Zhihua Fang, Liang He

Map-Free Visual Relocalization Enhanced by Instance Knowledge and Depth Knowledge 5639
Mingyu Xiao, Runze Chen, Haiyong Luo, Fang Zhao, Fan Wu, Hao Xiong, Xuepeng Ma, Juan Wang

Trusted Mamba Contrastive Network for Multi-View Clustering 5644
Jian Zhu, Xin Zou, Lei Liu, Zhangmin Huang, Ying Zhang, Chang Tang, Li-Rong Dai

Enhancing Listened Speech Decoding from EEG Via Parallel Phoneme Sequence Prediction 5649
Jihwan Lee, Tiantian Feng, Aditya Kommineni, Sudarsana Reddy Kadiri, Shrikanth Narayanan

Smr-Awarenet: An Adaptive Smr-Aware Neural Network for EEG Auditory Attention Guided Target Speech Extraction 5654
Xuefei Wang, Yuting Ding, Lei Wang, Fei Chen

SX-Stitch: An Efficient VMS-UNet Based Framework for Intraoperative Scoliosis X-Ray Image Stitching 5659
Yi Li, Heting Gao, Mingde He, Jinqian Liang, Jason Gu, Wei Liu

Ultra Lightweight Singing Melody Extraction Via Combination of Convolution and MLP 5664
Jun Liu, Kangjie Dong, Qiubo Huang, Shuai Yu, Wei Li

Implanting Robust Watermarks in Latent Diffusion Models for Video Generation 5669
Xiaohang Liu, Heng Chang, Jinfu Wei, Lei Zhu, Li Liu, Likun Li, Shiji Zhou, Chengyuan Li, Di Xu, Wei Gao

Near-Field FMCW SAR Imaging with Fast BPA 5674
Zhengguang Xu, Xiaoxu Liu

Difference Bonds Consistency and Complementarity to Enhance Multimodal Representation Learning 5679
Congbing He, Sensen Song, Zhenhong Jia, Hui Zhao

Enhancing Multilingual ASR for Unseen Languages Via Language Embedding Modeling 5684
Shao-Syuan Huang, Kuan-Po Huang, Andy T. Liu, Hung-Yi Lee

Black-Box Adversarial Defense Against Voice Conversion Using Latent Space Perturbation 5689
Jie Gao, Haiyun Li, Zhisheng Zhang, Zhiyong Wu

ZVEFusion: Zero-Shot Visual Enhancement Fusion for Infrared and Visible Images in Low Light 5694
Duo Liu, Yiqi Shi, Guoyin Zhang, Sizhao Li, Liguo Zhang

Robust Seizure Prediction Based on Riemannian Manifold Enhanced Denoising Adversarial Autoencoder 5699
Peizhen Peng, Yansheng Wu, Wanqi Yang, Wenxin Wei, Ming Yang

Generating is Believing: Membership Inference Attacks Against Retrieval-Augmented Generation 5704
Yuying Li, Gaoyang Liu, Chen Wang, Yang Yang

Training-Free Adapter for Multi-Modal Image Matching for All-Day Visual Place Recognition 5709
Anuradha Uggi, Sumohana Channappayya

Single Snapshot Direction of Arrival Estimation Using the EP-SURE-SBL Algorithm 5714
Fangqing Xiao, Dirk Slock

Quantum Run-Length Encoding: Optimizing Data Compression on Quantum Computers with Exponential Resource Efficiency 5719
Jiale Zhang, Xilong Che, Shiyong Jin, Kaifan Pan, Shun Peng, Juncheng Hu

An Attentive Dual-Encoder Framework Leveraging Multimodal Visual and Semantic Information for Automatic OSAHS Diagnosis 5724
Yingchen Wei, Xihe Qiu, Xiaoyu Tan, Jingjing Huang, Wei Chu, Yinghui Xu, Yuan Qi

RecNet: Optimization for Dense Object Detection in Retail Scenarios Based on View Rectification 5729
Junhao Xiao, Yi Chen, Xiao Feng, Ruoyu Wang, Zhiyu Wu

EgoNet: An Unified Egocentric Active Speaker Detection Framework for Both Camera Wearer and Visible Candidates 5734
Yongqian Li, Xin Zhou, Zheng He, Wei Yu, Yong Luo

RespDiff: An End-To-End Multi-Scale RNN Diffusion Model for Respiratory Waveform Estimation from PPG Signals 5739
Yuyang Miao, Zehua Chen, Chang Li, Danilo P. Mandic

A Proximal Variable Smoothing for Nonsmooth Minimization Involving Weakly Convex Composite with MIMO Application 5744
Keita Kume, Isao Yamada

CA-UAP: Content-Agnostic Universal Adversarial Perturbation for Enhanced Generalization 5749
Rui Lu, Ziqiang He, Jingyang Wen, Xiangui Kang, Z. Jane Wang

Event Masked Autoencoder: Point-Wise Action Recognition with Event-Based Cameras 5754
Jingkai Sun, Qiang Zhang, Jiaxu Wang, Jiahang Cao, Hao Cheng, Renjing Xu

FLAMO: An Open-Source Library for Frequency-Domain Differentiable Audio Processing 5759
Gloria Dal Santo, Gian Marco De Bortoli, Karolina Prawda, Sebastian J. Schlecht, Vesa Välimäki

Joint Training Framework for Accent and Speech Recognition Based on Conformer Low-Rank Adaptation 5764
Xuyi Zhuang, Yukun Qian, Shiyun Xu, Mingjiang Wang

Generalizable Articulated Object Perception with Superpoints 5769
Qiaojun Yu, Ce Hao, Xibin Yuan, Li Zhang, Liu Liu, Yukang Huo, Rohit Agarwal, Cewu Lu

Multi-Scale Parallel Hybrid Network for Palmprint Recognition 5774
Hao Yang, Shuyi Li, Bob Zhang, Yuqi Wang

NCDI-Diffusion: Neural Contextual and Directional Inversion for Novel View Synthesis Through Diffusion Models 5779
Wenpeng Xing, Jie Chen, Zaifeng Yang, Xin Tong, Changting Lin, Meng Han

Dual-Modality Guided Artistic Style Transfer with Pre-Trained Diffusion Models 5784
Jiaxiong Liu, Xiaolong Xiong, Jun Zhou

DN-DR: Discriminative Network with Dual Reconstruction for Image Anomaly Detection 5789
Wen Li, Chune Zhang

SFma-Unet: A Mamba-Based Spatial-Frequency Fusion Network for Medical Image Segmentation 5794
Zhanpeng Liu, Yuqiang Zhang, Bin Wang, Yang Yang, Lin Cai

HYB-VITON: A Hybrid Approach to Virtual Try-On Combining Explicit and Implicit Warping 5799
Kosuke Takemoto, Koshinaka Takafumi

Distribution Alignment Informed Thresholding for Semi-Supervised Curvilinear Structure Segmentation 5804
Yuhao Mo, Bo Peng, Bihan Wen, Xulei Yang, Ce Zhu, Xun Xu

Credible and Detailed 3D Face Reconstruction in Large Pose 5809
Xinyu Li, Xitie Zhang, Suping Wu, Ruijie Peng, Kehua Ma, Xiang Zhang

AMNS: Attention-Weighted Selective Mask and Noise Label Suppression for Text-To-Image Person Retrieval 5814
Runqing Zhang, Xue Zhou

Full-Text Error Correction for Chinese Speech Recognition with Large Language Model 5819
Zhiyuan Tang, Dong Wang, Shen Huang, Shidong Shang

MambaMOT: State-Space Model as Motion Predictor for Multi-Object Tracking 5824
Hsiang-Wei Huang, Cheng-Yen Yang, Wenhao Chai, Zhongyu Jiang, Jeng-Neng Hwang

GSMM: Efficient Global Sparsification for Resource-Conscious Multimodal Models 5829
Wenlun Zhang, Haoran Pang, Yucai Zhou, Shixiao Wang, Luking Li

A Conditional KAN Diffusion Network for Human Activity Recognition with Missing Sensor Signal Series 5834
Hao Xiong, Jiayi Gong, Haiyong Luo, Fang Zhao, Yang Gao, Runze Chen, Mingyu Xiao

MuRL-DTI: A Multimodal Feature Fusion Reinforcement Learning Approach for Cold Start in Drug-Target Interactions 5839
Yao Liu, Xin Wang, Ye Liu, Dandan Dou

ZOQO: Zero-Order Quantized Optimization 5844
Noga Bar, Raja Giryes

Learning in the Model Space: Fault Diagnosis by Co-Objective Learning in DynInt Model Space 5849
Ziyu Tang, Xiren Zhou, Shikang Liu, Chuyang Wei, Ao Chen, Huanhuan Chen

VulKiller: Java Web Vulnerability Detection with Code Property Graph and Large Language Models 5854
Xingchen Chen, Baizhu Wang, Mengjun Zhang, Yaqin Cao, Qixu Liu

SSAAD: A Multi-Scale Temporal-Frequency Graph Network for Binary Auditory Attention Detection with Self-Supervised Learning 5859
Shuai Huang, Yongxiong Wang, Huan Luo, Shuwen Jia, Han Chen, Chendong Qin, Zhongcai He, Rui Luo

When SparseMoE Meets Noisy Interactions: An Ensemble View on Denoising Recommendation 5864
Weipu Chen, Zhuangzhuang He, Fei Liu

Ultrasound-Guided Registration Pseudo-Labels for Semi-Supervised Brachial Plexus Segmentation 5869
Jia Luo, Ze Zhang, Yi Ding, Yiqian Wang, Jian Zhang

Swin-VasMamba: A Topologically Constrained Model for 3D Vascular Segmentation 5874
Ziyu Liu, Jiaxuan Li, Xiangjian He, Qing Xu, Xin Chen, Shoujun Zhou

CapT: A Hierarchical Capsule Representation Learning Approach for Class Continual Learning 5879
Ankita Chatterjee, Saransh Patel, Jayanta Mukhopadhyay, Partha Pratim Das

DPM-LVSN:A Diffusion Probabilistic Model-Based Left Ventricular Segmentation Network 5884
Jia Luo, Yuhao Zhong, Yi Ding, Chenghao Zhou, Minghui Pang

Commonality Augmented Disentanglement for Multimodal Crowdfunding Success Prediction 5889
Jiayang Li, Xovee Xu, Yili Li, Ting Zhong, Kunpeng Zhang, Fan Zhou

From Characters to Subwords: Modeling Unit Conversion for Low-Resource Speech Recognition 5894
Yizhi Wang, Haofei Zhang, Huiqiong Wang, Li Sun, Mingli Song

Object-Level Data Augmentation for Visual 3D Object Detection in Autonomous Driving 5899
Haoran Cheng, Junkai Xu, Liang Peng, Zheng Yang, Xiaofei He, Boxi Wu

Robust Target Speaker Direction of Arrival Estimation 5904
Zixuan Li, Shulin He, Xueliang Zhang

Enhancing Emotional Text-To-Speech Controllability with Natural Language Guidance Through Contrastive Learning and Diffusion Models 5909
Xin Jing, Kun Zhou, Andreas Triantafyllopoulos, Björn W. Schuller

Emotion-Preserving Prosody Anonymization Network for Voice Privacy Protection 5914
Jiabei He, Shiwan Zhao, Jiaming Zhou, Haoqin Sun, Hui Wang, Yong Qin

NTC-KWS: Noise-Aware CTC for Robust Keyword Spotting 5919
Yu Xi, Haoyu Li, Hao Li, Jiaqi Guo, Xu Li, Wen Ding, Kai Yu

MotionFlow: Joint Motion Priors and Appearance Enhancement for High-Accuracy Optical Flow Estimation 5924
Zixu Wang, Congxuan Zhang, Chen Zhen, Hongye Chen, Liyue Ge, Ke Lu

Universal Low-Resource Speech Synthesis Via Phoneme Fusion Coordinating Low-Rank Decomposition 5929
Yanliang Li, Zhengtao Yu, Linqin Wang, Shengxiang Gao, Ling Dong, Wenjun Wang

Flow-TSVAD: Target-Speaker Voice Activity Detection Via Latent Flow Matching for Speaker Diarization 5934
Zhengyang Chen, Bing Han, Shuai Wang, Yidi Jiang, Yanmin Qian

A Noisy Label Filter Based on GMM Binary Classification for Speaker Verification 5939
Jianglong Yao, Shenghui Lu, Pengyu Ren, Kaidi Wang, Qinyang Hong, Lin Li

NucleiFormer: A Nuclei Segmentation Model Optimized by Joint Haar Wavelet and Adaptive Feature Calibration 5944
Yulin Chen, Qian Huang, Zhijian Wang, Ziyang Yin, Meng Geng

HDRec: Hierarchical Distillation for Enhanced LLM-Based Recommendation Systems 5949
Lingyan Zhang, Wanyu Ling, Shuwen Daizhou, Li Kuang

Spatio-Temporal Mapping Generative Adversarial Network for Functional Connectivity Network Reconstruction Across Brain Atlases 5954
Hongzheng Guan, Tao Jin, Li Xiao, Gang Qu, Yu-Ping Wang

A Graph-Based Generative Adversarial Network Model for Inferring Task-State from Resting-State Functional Connectivity Networks 5959
Tao Jin, Hongzheng Guan, Li Xiao, Gang Qu, Yu-Ping Wang

DTR: Dynamic Tree-Ring Watermarking Framework for Diffusion-Based Video Generation 5964
Shunyang Zeng, Linlin Yang, Jin Yang, Yezhen Wang, Tianyu Gao

Achieving Robustness in Blind Modulo Analog-To-Digital Conversion 5969
Amir Weiss

AudioEditor: A Training-Free Diffusion-Based Audio Editing Framework 5974
Yuhang Jia, Yang Chen, Jinghua Zhao, Shiwan Zhao, Wenjia Zeng, Yong Chen, Yong Qin

Error Feedback Approach for Quantization Noise Reduction of Distributed Graph Filters 5979
Xue Xian Zheng, Tareq Al-Naffouri

Learning to Follow Infrared Prior Repersentation for Image Dehazing 5984
Yiquan Du, Ming Yang, Haiyang Huo, Jiaqi Shi, Jiangang Ding, Lili Pei

Frequency-Space Margin Perception for Open Set Knowledge Distillation 5989
Lijun Liu, Lihua Jing, Rui Wang, Yuan Wang, Zhishen Wang

Long-Tailed Oracle Character Recognition Based on Convolutional Neural Networks and Vision Transformers 5994
Zhongyuan Yang, Zhiwang Han, Alimjan Aysa, Ghalip Ibrahim, Kurban Ubul

Transferable Selective Virtual Sensing Active Noise Control Technique Based on Metric Learning 5999
Boxiang Wang, Dongyuan Shi, Zhengding Luo, Xiaoyi Shen, Junwei Ji, Woon-Seng Gan

Sensitivity of Room Impulse Responses in Changing Acoustic Environment 6004
Karolina Prawda

Soft Augmentation for Graph Classification 6009
Weihuang Zheng, Xiaotong Zhang, Rui Dong, YouYong Kong

Dynamic Sparse Encoding and Cross-Temporal Attention for Remote Sensing Image Change Detection 6014
Shaoxiong Lin, Tao Lei, Tongfei Liu, Shuxin Zhang, Chongdan Min, Asoke K. Nandi

Efficient Learning of Balanced Signed Graphs Via Iterative Linear Programming 6019
Haruki Yokota, Hiroshi Higashi, Yuichi Tanaka, Gene Cheung

Human Action Recognition in Multi-Level Convolutional Temporal Attention Network 6024
Qian Huang, ZhongQi Chen, Chang Li, WeiWen Qian

Reduced Spatial Dependency for More General Video-Level Deepfake Detection 6029
Beilin Chu, Xuan Xu, Yufei Zhang, Weike You, Linna Zhou

TDMER: A Task-Driven Method for Multimodal Emotion Recognition 6034
Qian Xu, Yu Gu, Chenyu Li, He Zhang, Hai-Xiang Lin, Linsong Liu

FASTER: Face Attribute Sliders with Semantic Rewards 6039
Jingyan Chen, Lanxiang Zhou, Han Fang, Zerun Feng, Chao Ban, Yaqi Li, Hao Sun, Jiani Hu

Image-Assisted Label Connective Completion for Vessel Segmentation with Insufficient Annotations 6044
Xiaoqi Zheng, Baoyao Yang, Xiuwen Fang, Wenfang Yao, Mang Ye

Flare-Aware RWKV for Flare Removal 6049
Wanying Zhang, Wei Shang, Dongwei Ren, Wangmeng Zuo

How Machines Perceive Rooms - Regions of Relevance in Room Impulse Responses 6054
Prachi Sharma, Christian Kehling

SAF: Local Shape-Aware Face-Based Garment Collision Handling Via Neural SDFs..... 6059
Minzhe Tang, Ruisheng Yuan, Dongliang Kou, Mingyang Sun, Lihua Zhang

Influence of Oropharyngeal Esophageal Cavity Geometry and Beak Angle on Vocal Tract Resonance of Birds Using Computational Modeling 6064
Noumida A, Rajeev Rajan

Deep Joint Source-Channel Coding for Wireless Point Cloud Transmission 6069
Cixiao Zhang, Mufan Liu, Wenjie Huang, Yin Xu, Yiling Xu, Dazhi He

Channel-Aware Domain-Adaptive Generative Adversarial Network for Robust Speech Recognition 6074
Chien-Chun Wang, Li-Wei Chen, Cheng-Kang Chou, Hung-Shin Lee, Berlin Chen, Hsin-Min Wang

Voxel-Sensitive Wavelet-Based Approach for Structural Distortion in Low-Dose CT Images..... 6079
Naragoni Saidulu, Vinit Kumar, Priya Ranjan Muduli

Self-Information Guided Speech Segmentation for Efficient Streaming ASR 6084
Wen Shen Teo, Yasuhiro Minami

Dual Position Attention Time-Frequency Network for Binaural Audio Synthesis 6089
Changjun He, Weiping Chen, Mingjiang Wang

Localised Frequency Latent Domain Watermarking of DDIM Generated Images..... 6094
Qiran Lai, Adrian G. Bors

Reducing the Gap Between Pretrained Speech Enhancement and Recognition Models Using a Real Speech-Trained Bridging Module 6099
Zhongjian Cui, Chenrui Cui, Tianrui Wang, Mengnan He, Hao Shi, Meng Ge, Caixia Gong, Longbiao Wang, Jianwu Dang

A Privacy-Preserving Cross-Modal Retrieval Scheme Based on CLIP and Deep Hashing..... 6104
Yuan Cao, Hui Zhang, Xinzheng Shang

Auxiliary Tasks Benefit Skeleton-Based Action Recognition 6109
Yuheng Yang, Haipeng Chen

Deeply Coupling EEG Signals and Eye Movements for Multi-Modal and Region-Aware Emotion Recognition 6114
Yuepeng Chen, Yue Gao, Xiaoling Fu, Hua He, Tianxiong Ouyang, Songling Chen, Xiangling Fu

SS-BRPE: Self-Supervised Blind Room Parameter Estimation Using Attention Mechanisms.....6119
Chunxi Wang, Maoshen Jia, Meiran Li, Changchun Bao, Wenyu Jin

Robust Supervised Graph Embedding Method for EEG-Based Brain Network Emotion Recognition 6124
Pengcheng Zhu, Cunbo Li, Peiyang Li, Fali Li, Dezhong Yao, Peng Xu

Benchmarking Music Generation Models and Metrics Via Human Preference Studies 6129
Florian Grötschla, Ahmet Solak, Luca A. Lanzendörfer, Roger Wattenhofer

CrossHash: Cross-Scale Vision Transformer Hashing for Image Retrieval 6134
Weigang Wang, Zhongwen Guo, Wenxiang Jiang, Yujun Lan, Wentao Ma

Indoor Airflow Imaging Using Physics-Informed Schlieren Tomography 6139
Arjun Teh, Wael H. Ali, Joshua Rapp, Hassan Mansour

F-StrIPE: Fast Structure-Informed Positional Encoding for Symbolic Music Generation 6144
Manvi Agarwal, Changhong Wang, Gaël Richard

GLST-GCN: Global-Local Spatio-Temporal Graph Convolutional Network for Skeleton-Based Hand Motion Prediction 6149
Wenrui Yang, Xinchun Yu, Xiao-Ping Zhang

WMAJL: Watcher-Mediated Attention Joint Learning Model for Multimodal Relation Extraction 6154
Yunrui Dong, Guiduo Duan, Tianxi Huang, Yunhao Li

LipReading for Low-Resource Languages by Language Dynamic LoRA 6159
Shuai Zou, Xuefeng Liang, Yiyang Huang

Lunar Tracking: A New Benchmark for Nighttime Tiny Object Tracking 6164
Mohammed Leo, Ding Zhang, Hai-Tao Zheng, Haiye Lin

Using Corrected ASR Projection to Improve AD Recognition Performance from Spontaneous Speech 6169
Yunfan Zhang, Yun Jin, Guanlin Chen, Yong Ma, Maoshen Jia, Peng Song

MDRNet: Multi-Branch with Different Feature Representations Network for Motor Imagery Classification 6174
Qiaoli Zhou, Yi Zhao, Shun Zhang, Jiawen Song, Qiang Du, Li Ke

LBPE: Long-Token-First Tokenization to Improve Large Language Models 6179
Haoran Lian, Yizhe Xiong, Zijia Lin, Jianwei Niu, Shasha Mo, Hui Chen, Peng Liu, Guiguang Ding

GaussianSlicer: Efficient Surface Reconstruction from Cross-Sectional Slices with Gaussian Splatting 6184
Yuhu Guo, Chenghao Qian, Yuhong Mo, Akkarit Sangpetch

Multi-Scale Feature Interaction and Adaptive Experts for Panoptic Segmentation in Remote Sensing Images 6189
Zhenkun Sun, Jia Liu, Wenhua Zhang, Fang Liu, Jingxiang Yang, Liang Xiao

SoundBeam Meets M2D: Target Sound Extraction with Audio Foundation Model 6194
Carlos Hernandez-Olivan, Marc Delcroix, Tsubasa Ochiai, Daisuke Niizumi, Naohiro Tawara, Tomohiro Nakatani, Shoko Araki

DEFormer: DCT-Driven Enhancement Transformer for Low-Light Image and Dark Vision 6199
Xiangchen Yin, Zhenda Yu, Xin Gao, Xiao Sun

SOA: A Sparsity-Oriented Activation on Sub-Layers of FFN of Transformers 6204
Yulong Meng, Yuan Li, Binhan Chen, Yi Kangt

IOVS4NeRF: Incremental Optimal View Selection for Large-Scale NeRFs 6209
Jingpeng Xie, Shiyu Tan, Yuanlei Wang, Tianle Du, Yifei Xue, Yizhen Lao

Resource Allocation for Semantic Segmentation Tasks in Autonomous Driving: A Likelihood Active Inference Approach 6214
Xinxin Zhu, F. Richard Yu, Ying He, Biao He

Stimulus Modality Matters: Impact of Perceptual Evaluations from Different Modalities on Speech Emotion Recognition System Performance 6219
Huang-Cheng Chou, Haibin Wu, Chi-Chun Lee

PFedFace: Personalized Federated Learning for Face Recognition 6224
Ya Gao, Anyang Su, Chenxu Zhao, Jie Song

Geometric Feature-Driven Metric Learning for 3D Craniofacial Superimposition 6229
Qingdong Long, Junli Zhao, Fuqing Duan, Chengyuan Wang, Xuesong Wang, Lijie Geng, Zhenkuan Pan, Mingquan Zhou

Same Semantics of the Signal - What Do We Cluster with What Representation 6234
Alexander Barnhill, Oliver Traub, Andreas Maier, Elmar Nöth, Christian Bergler

Carver: Learning to Reconstruct Right Ventricle from Sparse Multi-View 2D Echocardiograms 6239
Yida Li, Jun Shi, Zhaohui Wang, Tiantong Wang, Ziqi Zhu, Minfan Zhao, Junshi Chen, Hong An

Jointly Optimal Array Geometries and Waveforms in Active Sensing: New Insights into Array Design Via the Cramér-Rao Bound 6244
Ids van der Werf, Geert Leus, Robin Rajamäki

APTSniffer: Detecting APT Attack Traffic Using Retrieval-Augmented Large Language Models 6249
Hongbo Xu, Chengxiang Si, Zhouzhou, Chenxu Wang, Peishuai Sun, Qingyun Liu

MAFD: Fine-Grained Motion Style Transfer with Adaptive Signal Fusion 6254
Ziyun Qian, Dingkang Yang, Mingcheng Li, Dongliang Kou, Lihua Zhang

Privacy-Preserving Distributed Maximum Consensus Without Accuracy Loss 6259
Wenrui Yu, Richard Heusdens, Jun Pang, Qiongxiu Li

Bilevel Learning for Low-Light Image Enhancement and Detection 6264
Wei Wang, Bojian Song, Xi Chen, Yaxin Gao, Weimin Lei

Context-Aware Multi-Scale Polyp Segmentation Network 6269
Jinyao Zhou, Wenxin Yu, Zhiqiang Zhang, Jun Gong, Peng Chen, Chang Liu

A Label Co-Occurrence Transformation Network for Joint Empathy Detection and Empathy Intent Classification 6274
Liting Jiang, Di Wu, Haoxiang Su, Xiaoyong Guo, Shuangyong Song, Yanbing Li

Weakly-Supervised Learning Based Spine Instance Segmentation for MRI Planning 6279
Mu He, Xiaohan Hao, Lei Guo, Mengdie Song, Bensheng Qiu

A Generalized Graph Signal Processing Framework for Multiple Hypothesis Testing Over Networks 6284
Xingchao Jian, Martin Gölz, Feng Ji, Wee Peng Tay, Abdelhak M. Zoubir

MixSense : Mixture of Vision Sense 6289
Jian Lin, Zhuoran Wang, Qibo Qiu, Jianzhong Chen, Zixian Ge, Weizhong Jin, Yuchao Yan, Li Yu

LDGNet: LLMs Debate-Guided Network for Multimodal Sarcasm Detection 6294
Hengyang Zhou, Jinwu Yan, Yaqing Chen, Rongman Hong, Wenbo Zuo, Keyan Jin

An LSTM Feature Imitation Network for Hand Movement Recognition from sEMG Signals 6299
Chuheng Wu, S. Farokh Atashzar, Mohammad M. Ghassemi, Tuka Alhanai

Pancreatic Cystic Neoplasms Lesion Detection for Non-Contrast CT Image Via Teacher-Student Model 6304
Mengjie Pan, Qiu Guan, Zhiqiang Yang, Zhongwen Yu, Haixia Long, Xinli Xu, Ruihui Wang, Zhehao An, Feng Chen

Contactless Vital Sign Monitoring for Multiple People Using a Millimeter-Wave MIMO Radar 6309
Yuan Liu, Xuemei Fu, Ritesh Chandra Tewari, Xingze Wang, Andy W. H. Khong

From Voices to Beats: Enhancing Music Deepfake Detection by Identifying Forgeries in Background 6314
Zhaolin Wei, Dengpan Ye, Jiacheng Deng, Yuhan Lin

Diffusion Model with Multi-Layer Wavelet Transform for Low-Light Image Enhancement 6319
Haiyan Jin, Jing Wang, Fengyuan Zuo, Haonan Su, Zhaolin Xiao, Bin Wang, Yuanlin Zhang

SSCL-AMC: A Self-Supervised Automatic Modulation Classification Method Via Dynamic Augmentation and Ensemble Learning 6324
Yilin Cai, Dingzhao Li, Sheng Wu, Mingyuan Shao, Shaohua Hong, Haixin Sun

Audio Codec Augmentation for Robust Collaborative Watermarking of Speech Synthesis 6329
Lauri Juvela, Xin Wang

Improved Cross-Lingual Speaker Verification Using Speaker Sensitive Feature Guidance and Fine-Grained Phonetic Information 6334
Yongtai Ji, Guangxing Li, Hao Huang, Yanbing Li, Wushour Silamu

Dual Decoder for Fast Inference in Natural Language Generation 6339
Wenbo Wang, Huiying Wang, Zhaoyang Wang, Shuailou Li, Yu Wen

Enabling Beam Search for Language Model-Based Text-To-Speech Synthesis 6344
Zehai Tu, Guangyan Zhang, Yiting Lu, Adaeze Adigwe, Simon King, Yiwen Guo

MusicEval: A Generative Music Dataset with Expert Ratings for Automatic Text-To-Music Evaluation 6349
Cheng Liu, Hui Wang, Jinghua Zhao, Shiwan Zhao, Hui Bu, Xin Xu, Jiaming Zhou, Haoqin Sun, Yong Qin

Utterance as a Bridge: Few-Shot Joint Learning of Empathy Detection and Empathy Intent Classification 6354
Liting Jiang, Di Wu, Zhe Li, Shuangyong Song, Yanbing Li, Hao Huang

MSECG: Incorporating Mamba for Robust and Efficient ECG Super-Resolution 6359
Jie Lin, I Chiu, Kuan-Chen Wang, Kai-Chun Liu, Hsin-Min Wang, Ping-Cheng Yeh, Yu Tsao

OV-HHIR: Open Vocabulary Human Interaction Recognition Using Cross-Modal Integration of Large Language Models 6364
Lala Shakti Swarup Ray, Bo Zhou, Sungho Suh, Paul Lukowicz

ASCDomain: Domain Invariant Device-Adversarial Isotropic Knowledge Distillation Convolutional Neural Architecture 6369
Hubert Truchan, Tien Hung Ngo, Zahra Ahmadi

ChangeChat: An Interactive Model for Remote Sensing Change Analysis Via Multimodal Instruction Tuning 6374
Pei Deng, Wenqian Zhou, Hanlin Wu

Joint Semantic Segmentation of Optical and SAR Image in Hazy Environments Via Cross-Modal Information Rectification and Cross-Attention Fusion 6379
Xinyue Fan, Libao Zhang

Synergistic Integration of Cross-Spatial Learning for Lightweight Crack Detection 6384
Senyao Li, Jingling Yuan, Huilin Zhu, Xian Zhong

Text Enhancement Network for Complex Multi-Line Scene Text Image Super-Resolution 6389
Yang Liu, Yuliang Huang, Yiming Liu

Loss-Aware Curriculum Learning for Chinese Grammatical Error Correction 6394
Ding Zhang, Yangning Li, Lichen Bai, Hao Zhang, Yinghui Li, Haiye Lin, Hai-Tao Zheng, Xin Su, Zifei Shan

Adaptive Fine-Grained Feature Mining and RoI Feature Interaction Network for Small Object Detection in Aerial Images 6399
Fangmin Xie, Zhan Xiong, Zhuoyue Wang, Kun Yang, Guoqiang Xiao

BrainVis: Exploring the Bridge Between Brain and Visual Signals Via Image Reconstruction 6404
Honghao Fu, Hao Wang, Jing Jih Chin, Zhiqi Shen

Relation-Aware Semantic Alignment Network for Text-To-Image Person Retrieval 6409
Yong Wu, Rongxi Zhou, Hongchao Li, Ze Zhou, Feifei Wei, Min Li, Guodui He

PromptTA: Prompt-Driven Text Adapter for Source-Free Domain Generalization 6414
Haoran Zhang, Shuanghao Bai, Wanqi Zhou, Jingwen Fu, Badong Chen

LLFA: Fusing Global Illumination and Local Priors for Low-Light Face Image Enhancement with Adaptor 6419
Ziqian Shao, Tao Wang, Kaihao Zhang, Danhuai Zhao, Tong Lu

Enhancing Imaging Generation Through Implicit Neural Representations and HyperNetwork for Spatial Variability 6424
Jaehoon Cha, Siu-Lun Yeung, Siddharth Dhanpal, Jeyan Thiyagalingam

Which Prosodic Features Matter Most for Pragmatics? 6429
Nigel G. Ward, Divette Marco, Olac Fuentes

Cued Speech Generation Leveraging a Pre-Trained Audiovisual Text-To-Speech Model 6434
Sanjana Sankar, Martin Lenglet, Gérard Bailly, Denis Beautemps, Thomas Hueber

VP-NTK: Exploring the Benefits of Visual Prompting in Differentially Private Data Synthesis 6439
Chia-Yi Hsu, Jia-You Chen, Yu-Lin Tsai, Chih-Hsun Lin, Pin-Yu Chen, Chia-Mu Yu, Chun-Ying Huang

DOA Estimation of Coherent Sources Using Residual Network-Based Subspace Reconstruction 6444
Tiange Wang, Lingyu Chen, Huanglin Zhang

ROME: Radar Sparsity Improvement and Omnimodal Enhancement for 3D Object Detection in Bird's Eye Views 6449
Yilong Guo, Junyin Wang, Chenghu Du, Shengwu Xiong, Yaxiong Chen

Rethinking Camouflaged Object Detection Via Foreground-Background Interactive Learning 6454
Chenxi Zhang, Qing Zhang, Jiayun Wu

Robust and Efficient Text-Based Speech Editing Using Noise Conditioning and Rectified Flow 6459
Haowen Yin, Kai Wang, Hongli Yang, Hao Huang, Wushour Silamu

Estimation of Multi-Attribute Differential Graphs with Non-Convex Penalties 6464
Jitendra K. Tugnait

Knocking on IP: Unveiling Websites Through Cache-Aware Fingerprinting 6469
Yu Liu, Yifei Cheng, Yujia Zhu, Yong Ding, Yong Sun, Xiaoou Zhang

COCOLA: Coherence-Oriented Contrastive Learning of Musical Audio Representations 6474
Ruben Ciranni, Giorgio Mariani, Michele Mancusi, Emilian Postolache, Giorgio Fabbro, Emanuele Rodolà, Luca Cosmo

Char-SAM: Turning Segment Anything Model into Scene Text Segmentation Annotator with Character-Level Visual Prompts 6479
Enze Xie, Jiahao Lyu, Daiqing Wu, Huawen Shen, Yu Zhou

A Convolutional Recurrent Mixer Network for Radar Meteorological Image Super-Resolution 6484
Rafael Gonçalves Pires, Daniel F. S Santos, Roberto V. Calheiros, João Paulo Papa, Ik Hyun Lee, Sambit Bakshi, Khan Muhammad

HANet: A Harmonic Attention-Based Network for Singing Melody Extraction from Polyphonic Music 6489
Shijun Wang, Xiangzhu Kong, Hao Huang, Kai Wang, Ying Hu

Bridging Task Boundaries: Remote Sensing Image-Text Retrieval Via Dictionary-Driven Adaptation 6494
Junwei Xu, Tao Huang, Zhenyu Wang, Weisheng Dong, Xin Li

AGIAA-2K: A Fine-Grained Dataset for Aesthetic and Alignment Evaluation of AI-Generated Images 6499
Bo Hu, Nanxiang Li, Lihuo He, Wen Lu, Leida Li, Xinbo Gao

A Hierarchical Flow for Few-Shot Anomaly Detection Via Global-Local Aggregation Strategy 6504
Yan Wan, Tian Fan, Li Yao

Piano Transcription by Hierarchical Language Modeling with Pretrained Roll-Based Encoders 6509
Dichucheng Li, Yongyi Zang, Qiuqiang Kong

A Grouping Strategy-Based Progressive Fusion Network for Hyperspectral Image Super-Resolution 6514
Guohua Lv, Baodong Zhang, Yongbiao Gao, Guixin Zhao, Guotao Wang, Juncan Wang

EffectiveASR: A Single-Step Non-Autoregressive Mandarin Speech Recognition Architecture with High Accuracy and Inference Speed 6519
Ziyang Zhuang, Chenfeng Miao, Kun Zou, Ming Fang, Tao Wei, Zijian Li, Ning Cheng, Wei Hu, Shaojun Wang, Jing Xiao

A Ranking Scheme for Trust Region Multi-Agent Reinforcement Learning 6524
Ruichen Gao, Yi Hu, Deqin Zheng, Mengxuan Shao, Haiqi Zhu, Chenyue Song, Wei Zhang, Feng Jiang

A-PeARCNN: A Physics-Encoded AutoRegressive Convolutional Neural Network with AttentionNet for Solving Partial Differential Equations 6529
Yibo Han, Ruixuan Ren, Tiejun Li, Jingyi Chen, Jianmin Zhang

Resilient Test-Time Adaptation by Mitigating Batch-Normalization Overfitting 6534
Xingzhi Zhou, Boyang Zhang, Zhiliang Tian, Yibo Zhang, Xin Niu, Ka Chun Cheung, Simon See, Nevin L. Zhang

Pseudo-Labeling for Enhanced User Privacy in Approximate Machine Unlearning..... 6539
Shunichi Watanabe

LMAC-TD: Producing Time Domain Explanations for Audio Classifiers..... 6544
Eleonora Mancini, Francesco Paissan, Mirco Ravanelli, Cem Subakan

Enhancing Generalized EEG Classification with Decomposed Statistics-Diverse Feature Augmentation 6549
Yubin He, C.L. Philip Chen, Bianna Chen, Tong Zhang

Enhancing Multimodal Analogical Reasoning Through Triplet Interaction..... 6554
Hua Cai, Xuli Shen, Shuaishuai Li, Weilin Shen, Qing Xu

Audio Explanation Synthesis with Generative Foundation Models 6559
Alican Akman, Qiyang Sun, Björn W. Schuller

Semantic Consistency and Integrity Network for Cloth-Changing Person Re-Identification 6564
Anqi Wang, Liyan Zhang

GADACE: Graph Anomaly Detection Combining Attribute Contrast and Structure Reconstruction..... 6569
Shihui Wang, Yulan Yang, Zixin Tan, Zhigao Zheng, Jiawei Jiang, Hao Huang, Quanqing Xu, Chuanhui Yang

Comprehensive Feature Processing Based on Attention Mechanism for Co-Salient Object Detection 6574
Guohua Lv, Mao Yuan, Zengbin Zhang, Zhengyang Zhang, Zhenhui Ding, Guangxiao Ma

Towards Maximizing Semantic Coverage for Image-Text Retrieval 6579
Junhao Xu, Zheng Liu, Zhumin Chen, Fei Shen

D2S: Towards Efficient Sparse 3D Object Detection Via Dense to Sparse Knowledge Distillation..... 6584
Yuqi Huang, Longjun Liu, Yingke Gao, Haonan Zhang, Haoteng Li

Phone-Purity Guided Discrete Tokens for Dysarthric Speech Recognition 6589
Huimeng Wang, Xurong Xie, Mengzhe Geng, Shujie Hu, Haoning Xu, Youjun Chen, Zhaoqing Li, Jiajun Deng, Xunying Liu

DiffDesign: A Diffusion Model Using Garment Knowledge-Enhanced for Fashion Design Synthesis..... 6594
Shouhao Wu

SYKI-SVC: Advancing Singing Voice Conversion with Post-Processing Innovations and an Open-Source Professional Testset 6599
Yiquan Zhou, Wenyu Wang, Hongwu Ding, Jiacheng Xu, Jihua Zhu, Xin Gao, Shihao Li

USV-AUV Collaboration Framework for Underwater Tasks Under Extreme Sea Conditions..... 6604
Jingzehua Xu, Guanwen Xie, Xinqi Wang, Yimian Ding, Shuai Zhang

CGNet: Classification-Guided Multi-Task Interactive Network for Hyperspectral and Multispectral Image Fusion 6609
Guohua Lv, Yanlong Xu, Yongbiao Gao, Guixin Zhao, Guotao Wang, Xiangcheng Sun

Multilevel Semantic-Aware Model for AI-Generated Video Quality Assessment 6614
Jiaze Li, Haoran Xu, Shiding Zhu, Junwei He, Haozhao Wang

EfficientSleepNet: A Novel Lightweight End-To-End Model for Automated Sleep Staging on Single-Channel EEG 6619
Fei Wang, Zekun Zheng, BangShun Hu, Xiaodong Yang, Maolin Tang, Haiyun Huang

Aligning Text-To-Image Diffusion Models Without Human Feedback 6624
Tao Liu, Huafeng Kuang, Xianming Lin

Effective and Efficient Mixed Precision Quantization of Speech Foundation Models 6629
Haoning Xu, Zhaoqing Li, Zengrui Jin, Huimeng Wang, Youjun Chen, Guinan Li, Mengzhe Geng, Shujie Hu, Jiajun Deng, Xunying Liu

Autoregressive Density Estimation Transformers for Multivariate Time Series Anomaly Detection 6634
Mohammed Ayalew Belay, Adil Rasheed, Pierluigi Salvo Rossi

One-Shot Talking Face Generation with Expression Editing 6639
Ganyu Huang, Liping Shen

The Impact of Decorrelation on Transformer Interpretation Methods: Applications to Clinical Speech AI 6644
Lingfeng Xu, Kimberly D. Mueller, Julie Liss, Visar Berisha

GCAT: Gated Convolutional Attention Transformer for Efficient Image Super-Resolution 6649
Zhijian Wu, Kaiyi Feng, Dingjiang Huang

Iterative Operator Sketching Framework for Large-Scale Imaging Inverse Problems 6654
Junqi Tang, Subhadip Mukherjee, Carola-Bibiane Schönlieb

EvaSR: Rethinking Efficient Visual Attention Design for Image Super-Resolution 6659
Zhijian Wu, Chenhan Zhang, Dingjiang Huang

Cross-Domain Few-Shot Open-Set Keyword Spotting Using Keyword Adaptation and Prototype Reprojection 6664
Mingru Yang, Qianhua He, Jinxin Huang, Yongqiang Chen, Zunxian Liu, Yanxiong Li

Blind Spatial Impulse Response Generation from Separate Room- And Scene-Specific Information 6669
Francesc Lluís, Nils Meyer-Kahlen

Information-Theoretic Minimax Regret Bounds for Reinforcement Learning Based on Duality 6674
Raghav Bongole, Amaury Gouverneur, Borja Rodríguez-Gálvez, Tobias J. Oechtering, Mikael Skoglund

Structure-Preserving Video Hashing Via Self-Supervised Transformer for Retrieval 6679
Lixia Du, Xiaoping Liang, Likai Yang, Zhenjun Tang

Low-Resolution Hierarchical Training for Efficient 3D Gaussian Splatting 6684
Yilin Jin, Shaohui Li, Zhi Li, Yu Liu

Graph Neural Networks for Parkinson's Disease Detection 6689
Shakeel A. Sheikh, Yacouba Kaloga, Md. Sahidullah, Ina Kodrasi

Improving Food Recognition with Retrieval-Augmented and Domain-Adaptive LVLMs 6694
Dehua Ma, Zhenbo Xu, Tianshun Xing, Lu Yuan, Jinghan Yang, Huijia Wu, Ming Lei, Zhaofeng He

HCoTT: Hierarchical Chain-Of-Thought Distillation 6699
Zhichang Wang, Xianwei Zhuang, Zhihong Zhu, Yuexian Zou

Prompt Fusion and Aspect-Oriented Filtration for Aspect-Based Multimodal Sentiment Analysis 6704
Jiachang Sun, Xiuhong Li, Fuxian Zhu

Differentially Private Distribution Estimation Using Functional Approximation 6709
Ye Tao, Anand D. Sarwate

Annealing Distillation Algorithm for Transferring Unsupervised Clustering Knowledge to Supervised Student Models 6714
Jiaying Gao, Fausto Giunchiglia, Tongyu Zhao, Hao Xu

Positioning and Transmission in Cell-Free Networks: Ambiguity Function, and MRC/MRT Array Gains..... 6719
Luc Vandendorpe, Laurence Defraigne, Guillaume Thiran, Thomas Pairon, Christophe Craeye

Hyperbolic Distance Based on EMD and Diffusion for Hyperspectral Imaging..... 6724
Elad Lavi, Amir Bourvine, Ya-Wei Eileen Lin, Ronen Talmon

LTOS: Layout-Controllable Text-Object Synthesis Via Adaptive Cross-Attention Fusions 6729
Xiaoran Zhao, Tianhao Wu, Yu Lai, Zhiliang Tian, Zhen Huang, Yahui Liu, Zejiang He, Dongsheng Li

RAOCSL: A BERT-Based Strategy for Identifying Learner Confusion Under Class Imbalance..... 6734
Tongyu Zhao, Jiaying Gao, Yu Feng, Yatong Zu, Adriano Tavares, Tiago Gomes, Sandro Pinto, Hao Xu

A2GP-SF: Enhancing Few-Shot Class Incremental Learning Via Attribute Generative Prompting and Adaptive Sharpness Flattening..... 6739
Zhiming Chen, Desen Wang, Sisi Fu, Congcong Wen, Hui Lin, Bingzhi Chen

SVRM: Composing Various Network Service Fuzzing Corpus with One Single Model..... 6744
Wenfeng Lin, Zhiyuan Jiang, Fangliang Xu, Yunfei Su, Zhiwei Li, Lingchu Mao, Chaojing Tang

Let There Be Light: Robust Lensless Imaging Under External Illumination with Deep Learning 6749
Eric Bezzam, Stefan Peters, Martin Vetterli

SINET: Sparsity-Driven Interpretable Neural Network for Underwater Image Enhancement..... 6754
Gargi Panda, Soumitra Kundu, Saumik Bhattacharya, Aurobinda Routray

Unlocking Financial Statement Fraud Detection: Tracking Disclosure Changes Via Representation Learning 6759
Yue Yu, Zhen Wu, Yanni Han, Zhuoqun Li, Wenqi Wei

Modifying Flow Matching for Generative Speech Enhancement 6764
Roman Korostik, Rauf Nasretdinov, Ante Jukić

Advancing Few-Shot Class-Incremental Learning with Virtual Prototype Guidance Prompting..... 6769
Xiang Qiu, Huanjia Zhu, Xiaocheng Fang, Jun Liang, Bingzhi Chen, Hui Lin

SPSinger: Multi-Singer Singing Voice Synthesis with Short Reference Prompt 6774
Junchuan Zhao, Chetwin Low, Ye Wang

Spherical Sector Harmonics Based Acoustic Source Localization Using Circular Array 6779
Deepika Kumari, Lalan Kumar

Speech Separation for Low-Resource Languages 6784
Marvin Borsdorf, Zexu Pan, Pascal Himmelmann, Haizhou Li, Tanja Schultz

CHASE: Channel-Wise and Spatial Attention for Early Exiting in Image Classification 6789
Youva Addad, Alexis Lechervy, Frédéric Jurie

Leveraging Joint Spectral and Spatial Learning with MAMBA for Multichannel Speech Enhancement 6794
Wenze Ren, Haibin Wu, Yi-Cheng Lin, Xuanjun Chen, Rong Chao, Kuo-Hsuan Hung, You-Jin Li, Wen-Yuan Ting, Hsin-Min Wang, Yu Tsao

VoicePrompter: Robust Zero-Shot Voice Conversion with Voice Prompt and Conditional Flow Matching 6799
Ha-Yeong Choi, Jaehan Park

DARN: An Attention-Based Neural Network Using Residual Blocks for Sleep Micro-Events Detection 6804
Fei Wang, Zhuorong Li, Jingcong Li

BAD: Bidirectional Auto-Regressive Diffusion for Text-To-Motion Generation 6809
Seyed Rohollah Hosseyni, Ali Ahmad Rahmani, Seyed Jamal Seyedmohammadi, Sanaz Seyedin, Arash Mohammadi

Uncertainty-Aware Masked Modeling in Medical Imaging 6814
Jiayu Zhang, Yuxin Cao, Dexuan Xu

Particle-Based Data-Driven Nonlinear State Estimation of Model-Free Process from Nonlinear Measurements 6819
Anubhab Ghosh, Yonina C. Eldar, Saikat Chatterjee

Dynamic ROI Adaptation for Accurate Non-Contact Heart Rate Estimation Using VGG-13 Based Encoder-Decoder Model and Facial Landmarks 6824
Aravind A Anil, Srinivasa Karthik, Mohanasankar Sivaprakasam, Jayaraj Joseph

Multi-Source Multi-Target Domain Similarity Network for Cross-Cultural EEG Emotion Recognition 6829
Haiqing Hu, Hanwen Shi, Bao-Liang Lu, Wei-Long Zheng

A Multi-Prior Fusion Network for Video-Based Micro-Expression Recognition 6834
Chuang Ma, Shaokai Zhao, Yu Pei, Liang Xie, Erwei Yin, Ye Yan

Automotive Radar Target Detection in Widely Separated and Distributed Aperture Radar Systems 6839
Moein Ahmadi, Björn Ottersten, Bhavani Shankar M. R., Thomas Stifter

Sequential Diffusion-Guided Deep Image Prior for Medical Image Reconstruction 6844
Shijun Liang, Ismail Alkhouri, Qing Qu, Rongrong Wang, Saiprasad Ravishankar

Lightweight Clustered Federated Learning Via Feature Extraction 6849
Guanzhang Lao, Xinglin Zhang, Yun Li, Yue-Jiao Gong

TSLA: A Multi-Task Time Series Language Model 6854
Liri Fang, Yuncong Chen, Wenchao Yu, Yanchi Liu, Lu-an Tang, Vetle I. Torvik, Haifeng Chen

Efficient Gridless Wideband Direction-Of-Arrival Estimation from Many Frequencies 6859
Yiming Zhou, Huayu Fu, Wei Dai

IDANSE: Iterative Data-Driven Nonlinear State Estimation of Model-Free Hidden Sequences 6864
Hang Qin, Anubhab Ghosh, Saikat Chatterjee

Classifier-Guided Captioning Across Modalities 6869
Ariel Shaulov, Tal Shaharabany, Eitan Shaar, Gal Chechik, Lior Wolf

Conformal Prediction for Manifold-Based Source Localization with Gaussian Processes 6874
Vadim Rozenfeld, Bracha Laufer Goldshtein

DocVideoQA: Towards Comprehensive Understanding of Document-Centric Videos Through Question Answering 6879
Haochen Wang, Kai Hu, Liangcai Gao

FreeSVC: Towards Zero-Shot Multilingual Singing Voice Conversion 6884
Alef Iury Ferreira, Lucas Rafael Gris, Augusto Da Rosa, Frederico Oliveira, Edresson Casanova, Rafael Sousa, Arnaldo Junior, Anderson Soares, Arlindo Galvão Filho

Speech Enhancement with MAP-Based Training for Robust ASR 6889
You-Jin Li, Rong Chao, Borching Su, Yu Tsao

Right Label Context in End-To-End Training of Time-Synchronous ASR Models 6894
Tina Raissi, Ralf Schlüter, Hermann Ney

Digital Operating Mode Classification of Real-World Amateur Radio Transmissions 6899
Maximilian Bundscherer, Thomas Schmitt, Ilja Baumann, Tobias Bocklet

Extremum Encoding for Joint Baseband Signal Compression and Time-Delay Estimation for Distributed Systems 6904
Amir Weiss, Yuval Kochman, Gregory W. Wornell

Low Complexity Rate Splitting Approach in RIS-Aided Systems Based on Channel Statistics 6909
Sadaf Syed, Michael Joham, Wolfgang Utschick

RaLU-Net: Deep Unfolded Radar Localization of Humans for Precise Multi-Person Non-Contact Vital Signs Monitoring 6914
Yonathan Eder, Yhonatan Kvich, Rui Guo, Yonina C. Eldar

Threshold Sensitivity in Two-Channel Modulo ADCs: Analysis and Robust Reconstruction 6919
Wenyi Yan, Lu Gan, Yimin D. Zhang

Hierarchical Bayesian Estimation of COVID-19 Reproduction Number 6924
Patrice Abry, Juliette Chevallier, Gersende Fort, Barbara Pascal

Text-Guided Device-Realistic Sound Generation for Fiber-Based Sound Event Classification 6929
Wataru Kohno, Noriyuki Tonami, Jian Fang, Shaobo Han, Jingchen Sun, Ting Wang

ATGnet: Adaptive Temporal Graph Network for EEG-Enabled Sound Source Tracking in Cocktail Party Scenarios 6934
Saurav Pahuja, Gabriel Ivucic, Siqi Cai, Dashanka De Silva, Tanja Schultz, Haizhou Li

RTF Estimation Using Riemannian Geometry for Speech Enhancement in the Presence of Interferences 6939
Or Ronai, Yuval Sitton, Amitay Bar, Ronen Talmon

Can Large Audio-Language Models Truly Hear? Tackling Hallucinations with Multi-Task Assessment and Stepwise Audio Reasoning 6944
Chun-Yi Kuan, Hung-Yi Lee

Revisiting Neighborhood Aggregation in Graph Neural Networks for Node Classification Using Statistical Signal Processing 6949
Mounir Ghogho

Username-Password Models Beyond Traditional Password Guessability Assessment 6954
Jiahong Yang, Wenting Li, Haibo Cheng, Ping Wang

Explainable Reinforcement Learning for Trajectory Design in UAV-Assisted Wireless Networks 6959
Ali Krayani, Khalid Khan, Lucio Marcenaro, Carlo Regazzoni

Intelligent Target Maneuverability in Presence of Tracking with Multiple Radars 6964
Bhavani Shankar Mysore Rama Rao, Jyoti Bhatia, Kunwar Pritiraj Rajput, Bjorn Ottersten

Convolutional Sparse Coding with Multipath Orthogonal Matching Pursuit 6969
Yanis Gomes, Charles Truong, Jean-Philippe Saut, Fikri Hafid, Pascale Prieur, Laurent Oudre

Guitar-TECHS: An Electric Guitar Dataset Covering Techniques, Musical Excerpts, Chords and Scales Using a Diverse Array of Hardware 6974
Hegel Pedroza, Wallace Abreu, Ryan M. Corey, Iran R. Roman

A Counterfactual Ultrasound Anti-Interference Self-Supervised Network for B-Mode Ultrasound Tongue Extraction 6979
Yan Jia, Yuqing Cheng, Kele Xu, Yong Dou, Peng Qiao, Zhouyu He

DEGSTalk: Decomposed Per-Embedding Gaussian Fields for Hair-Preserving Talking Face Synthesis 6984
Kaijun Deng, Dezhi Zheng, Jindong Xie, Jinbao Wang, Weicheng Xie, Linlin Shen, Siyang Song

Infant Cry Detection Using Causal Temporal Representation 6989
Minghao Fu, Danning Li, Aryan Gadhiya, Benjamin Lambright, Mohamed Alowais, Mohab Bahnassy, Saad El Dine Elletter, Hawau Olamide Toyin, Haiyan Jiang, Kun Zhang, Hanan Aldarmaki

Learning to Optimally Sample in MRI for Denoising-Driven Regularization 6994
Pavan Kumar Reddy K, Kunal N. Chaudhury

SML: A Backdoor Defense for Non-Intrusive Speech Quality Assessment Via Semi-Supervised and Multi-Task Learning 6999
Ying Ren, Wenjie Zhang, Jiahong Ye, Jie Li, Diqun Yan, Bin Ma

Automated Extraction of Spatio-Semantic Graphs for Identifying Cognitive Impairment 7004
Si-Ioi Ng, Pranav S. Ambadi, Kimberly D. Mueller, Julie Liss, Visar Berisha

Generalized Linear Models with 1-Bit Measurements: Asymptotics of the Maximum Likelihood Estimator 7009
Jaimin Shah, Martina Cardone, Cynthia Rush, Alex Dytso

AC-Mix: Self-Supervised Adaptation for Low-Resource Automatic Speech Recognition Using Agnostic Contrastive Mixup 7014
Carlos Carvalho, Alberto Abad

Schoenberg Kernel Loss for Spiking Neural Network Training 7019
Liyang Ru, Kan Li, José C. Príncipe

Multichannel Modulo Sampling with Unlimited Noise 7024
Dorian Florescu

LMTalker: Sparse Landmark-Guided Gaussian Splatting for High-Fidelity Talking Head Synthesis 7029
Zhifeng Xie, Zhiwen Jiang, Xuemin Lei, Mengtian Li

Federated Learning with Heterogeneous Feature Adaptation for Human Activity Recognition 7034
Jiaqi Wang, Tobias Schlagenhauf, Setareh Maghsudi

A Study of Improving the Privacy-Utility Trade-Off of Task-Specific Models with Learnable Privacy 7039
Savas Ozkan, Taha Ceritli, Jeongwon Min, Eunchung Noh, Jung Min Cho, Dookun Park, Mete Ozay

Colorization Network Watermarking in the CIE-Lab Domain 7044
Alessio Chiovelli, Nischay Purnekar, Benedetta Tondi, Mauro Barni

Rate-Constrained Quantization for Communication-Efficient Federated Learning 7049
Shayan Mohajer Hamidi, Ali Bereyhi

Multimodal Emotion Recognition in Conversation Via Possible Speaker's Audio and Visual Sequence Selection 7054
Rahul Singh Maharjan, Niyati Rawal, Marta Romeo, Lorenzo Baraldi, Rita Cucchiara, Angelo Cangelosi

Hyper-Refinement for Low-Rank Adaptation 7059
Savas Ozkan, Taha Ceritli, Jeongwon Min, Eunchung Noh, Jung Min Cho, Dookun Park, Mete Ozay

Multiclass Arrhythmia Classification Using Smartwatch Photoplethysmography Signals Collected in Real-Life Settings 7064
Dong Han, Jihye Moon, Luís Roberto Mercado Díaz, Darren Chen, Devan Williams, Eric Y. Ding, Khanh-Van Tran, David D. McManus, Ki H. Chon

KAN V.s. MLP for Offline Reinforcement Learning 7069
Haihong Guo, Fengxin Li, Jiao Li, Hongyan Liu

Music2Latent2: Audio Compression with Summary Embeddings and Autoregressive Decoding 7074
Marco Pasini, Stefan Lattner, György Fazekas

Quasi Polynomial and Interpolative Models for Tensor Approximation 7079
Jeongmin Chae, Selin Bac, Usama Saleem, Shaama Mallikarjun Sharada, Urbashi Mitra

RIS-Enabled Self-Interference Elimination in Monostatic Full-Duplex DFRC Systems 7084
Linlong Wu, Zichao Xiao, Bhavani Shankar M. R., Björn Ottersten

Generative Diffusion Model-Based Energy Management in Networked Energy Systems 7089
Xinyu Lu, Zhanbo Feng, Jiawei Sun, Jiong Lou, Chentao Wu, Wugedele Bao, Jie Li

Regarding the Existence of the Internal Language Model in CTC-Based E2E ASR 7094
Zeyu Zhao, Peter Bell

Compressive Imaging Reconstruction Via Conditional Diffusion Model with Augmented Measurements..... 7099
Emmanuel Martinez, Leon Suarez, Romario Gualdrón-Hurtado, Roman Jacome, Henry Arguello

Cauchy-Schwarz Divergence Transfer Entropy 7104
Zhaozhao Ma, Shujian Yu

Unsupervised Domain Adaptation Via Data Pruning 7109
Andrea Napoli, Paul White

Dysarthric Speech Conformer: Adaptation for Sequence-To-Sequence Dysarthric Speech Recognition 7114
Qianli Wang, Zihan Zhong, Satwinder Singh, Clarion Mendes, Mark Hasegawa-Johnson, Waleed Abdulla, Seyed Reza Shahamiri

Point-UMAE: Unet-Like Masked Autoencoders for Point Cloud Self-Supervised Learning 7119
Hongliang Zeng, Ping Zhang, Fang Li, Tingyu Ye, Jiahua Wang, Xianbo Yang

Chinese Speech Processing Via Chinese Character Feature..... 7124
Rui Jiang, Zhao Yang, Wei Xi, Xiao Fu, Jizhong Zhao

Track-MDP: Reinforcement Learning for Target Tracking with Controlled Sensing..... 7129
Adarsh M. Subramaniam, Argyrios Gerogiannis, James Z. Hare, Venugopal V. Veeravalli

Targeted Data Poisoning for Black-Box Audio Datasets Ownership Verification..... 7134
Wassim Wes Bouaziz, El-Mahdi El-Mhamdi, Nicolas Usunier

Text Descriptions of Actions and Objects Improve Action Anticipation..... 7139
Apoorva Beedu, Harish Haresamudram, Irfan Essa

L2G: Head Gesture Animation Using an Emotion Guided Language Model 7144
Rishabh Agrawal, Apurva Narayan

BloomCoreset: Fast Coreset Sampling Using Bloom Filters for Fine-Grained Self-Supervised Learning 7149
Prajwal Singh, Gautam Vashishtha, Indra Deep Mastan, Shanmuganathan Raman

Estimating the Number and Locations of Boundaries in Reverberant Environments with Deep Learning 7154
Toros Arikan, Luca M. Chackalackal, Fatima Ahsan, Konrad Tittel, Andrew C. Singer, Gregory W. Wornell, Richard G. Baraniuk

Mamba Fusion: Learning Actions Through Questioning 7159
Apoorva Beedu, Zhikang Dong, Jason Sheinkopf, Irfan Essa

Low-Rank Tensors for Multi-Dimensional Markov Models..... 7164
Madeline Navarro, Sergio Rozada, Antonio G. Marques, Santiago Segarra

Massive MIMO Channel-Aware Decision Fusion Aided by Reconfigurable Intelligent Surfaces..... 7169
Domenico Ciuonzo, Alessio Zappone, Marco Di Renzo, Linlong Wu

Generative Sensing: Pre-Training LiDAR with Masked Autoencoders for Ultra-Frugal Perception..... 7174
Sina Tayebati, Theja Tulabandhula, Amit Ranjan Trivedi

Elevating Robust ASR by Decoupling Multi-Channel Speaker Separation and Speech Recognition 7179
Yufeng Yang, Hassan Taherian, Vahid Ahmadi Kalkhorani, DeLiang Wang

Self-Prompting Polyp Segmentation in Colonoscopy Using Hybrid YOLO-SAM2 Model 7184
Mobina Mansoori, Sajjad Shahabodini, Jamshid Abouei, Konstantinos N. Plataniotis, Arash Mohammadi

Diffusion-Based Unsupervised Audio-Visual Speech Enhancement 7189
Jean-Eudes Ayilo, Mostafa Sadeghi, Romain Serizel, Xavier Alameda-Pineda

K-HashFed: Communication Efficient Federated Learning Through Gradient Clustering and Hashing 7194
Ayshika Kapoor, Dheeraj Kumar

Fair CoVariance Neural Networks 7199
Andrea Cavallo, Madeline Navarro, Santiago Segarra, Elvin Isufi

Multi-Channel Speaker Counting for EEND-VC-Based Speaker Diarization on Multi-Domain Conversation 7204
Naohiro Tawara, Atsushi Ando, Shota Horiguchi, Marc Delcroix

Correlated Attention in Transformers for Multivariate Time Series 7209
Quang Minh Nguyen, Lam M. Nguyen, Subhro Das

Multi-Scale Spectral Mixture Neural Operator 7214
Fengrui Jing, Hongzhen Ding, TaoSong

SCS: Spatially Consistent Self-Supervised Approach for One-Shot Anatomical Landmark Detection 7219
Lu Han, Boyu Chen, Zherui Zhang, Li Guo, Shibiao Xu

MS-SCANet: A Multiscale Transformer-Based Architecture with Dual Attention for No-Reference Image Quality Assessment 7224
Mayesha Maliha R. Mithila, Mylène C. Q. Farias

Hard Sample Aware Robust Contrastive Learning for Multi-View Clustering 7229
Yuanzhe Cai, Zhikui Chen, Jing Gao, Peng Li, Jianing Zhang

Phoneme-Aware Acoustic Analysis of Natural Speech for Lung Function Assessment 7234
Sejal Bhalla, Tien Han, Andrea Gershon, Robert Wu, Eyal de Lara, Alex Mariakakis

INFR-GC: Interpretable Feature Representations for Granger Causality in Cortico-Muscular Interactions 7239
Farwa Abbas, Verity McClelland, Wei Dai, Zoran Cvetkovic

Lossless Phase Conversion Method for Object Wave-Based Hologram Compression 7244
Hiroki Kojima, Ryota Koiso, Ryosuke Watanabe, Keisuke Nonaka

Trainingless Adaptation of Pretrained Models for Environmental Sound Classification 7249
Noriyuki Tonami, Wataru Kohno, Keisuke Imoto, Yoshiyuki Yajima, Sakiko Mishima, Reishi Kondo, Tomoyuki Hino

Multi-Task Learning for Ultrasonic Echo-Based Depth Estimation with Audible Frequency Recovery 7254
Junpei Honma, Akisato Kimura, Go Irie

On the Role of Prompt Construction in Enhancing Efficacy and Efficiency of LLM-Based Tabular Data Generation 7259
Banooqa Banday, Kowshik Thopalli, Tanzima Z. Islam, Jayaraman J. Thiagarajan

A Novel Compressive Compound Word Encoding and Independent Word Attention for Symbolic Music Generation 7264
Lianyu Zhou, Liang Yin, Yukun Qian

I-KAN: Reconstructing Over-Range Inertial Signals 7269
Yifeng Wang, Shu Zhang, Yi Zhao

Multi-Head Auto-Correlation Attention Networks for Session-Based Social Recommendation 7274
Mengying Lu, Xingyu Lu, Hai-Tao Zheng, Zhao Wei, Yong Xu, Bingxu An

LkSFocalNets: Video Action Recognition with Large Kernel Selective Focal Networks 7279
Jian Xiao, Ping Shi, Qipei Li

Optimizing Biomarkers from Earbud Ballistocardiogram: Calibration and Calibration-Free Algorithms for Accelerometer Axis Selection and Fusion 7284
Yunzhi Li, Md Mahbubur Rahman, Mehrab Bin Morshed, Md Saiful Islam, Hao Zhou, Weinan Wang, Holland Ernst, Li Zhu, Jilong Kuang

Tessellated Linear Model for Age Prediction from Voice 7289
Dareen Alharthi, Mahsa Zamani, Bhiksha Raj, Rita Singh

Enhancing 3D Medical Image Understanding with 2D Multimodal Large Language Models 7294
Qiuhui Chen, Xuancheng Yao, Huping Ye, Yi Hong

AudioBERT: Audio Knowledge Augmented Language Model 7299
Hyunjong Ok, Suho Yoo, Jaeho Lee

A Domain Adversarial Learning Framework for Major Depression Disorder Diagnosis 7304
Shaozhe Liu, Leike An, Ziyu Jia

Foundation Model and Temporal Priors-Guided Transductive Few-Shot Action Recognition 7309
Bach Vu, Hoang Nguyen, Quang Minh Nguyen, Duong Le, Hieu Pham, Phi Le Nguyen, Lam M. Nguyen

Data-Free Post-Training Quantization with Block-Wise Enhanced Sample Generation 7314
Ruiyao Zhang, Zhiwei Dong, Long Huang, Shutong Ti, Songlu Chen, Xucheng Yin

EGENN: An Efficient Graph-Enhanced Neural Network for Multivariate Time Series Forecasting 7319
Haoxuan Xu, Haiqi Zhu, Yifan Chen, Chunzhi Yi, Baichun Wei, Feng Jiang

Apollo: Band-Sequence Modeling for High-Quality Audio Restoration 7324
Kai Li, Yi Luo

Preventing Output Saturation in Active Noise Control: An Output-Constrained Kalman Filter Approach 7329
Junwei Ji, Dongyuan Shi, Boxiang Wang, Xiaoyi Shen, Zhengding Luo, Woon-Seng Gan

Sub-Band Domain Multi-Hypothesis Acoustic Echo Canceler Based Acoustic Scene Analysis 7334
Benjamin J. Southwell, Yin-Lee Ho, David Gunawan

Large Covariance Matrix Estimation for Groups of Highly Correlated Variables Via Nonconvex Optimization 7339
Shanshan Zou, Ziping Zhao

Fast Word Error Rate Estimation Using Self-Supervised Representations for Speech and Text 7344
Chanho Park, Chengsong Lu, Mingjie Chen, Thomas Hain

HyperSF: A Hypergraph Representation Learning Method Based on Structural Fusion 7349
Xiangfei Fang, Chengying Huan, Boying Wang, Shaonan Ma, Heng Zhang, Chen Zhao

Adaptive Contribution Modulation for Multi-Modal Manipulation Media Detection and Grounding 7354
Yixiang Li, Biao Leng

A Framework Based on Data Augmentation for Knowledge Graph Entity Typing 7359
Zepeng Li, Minyu Zhai, Xiao Li, Rikui Huang, Chenhui Liang, Bin Hu

HyperKAN: Hypergraph Representation Learning with Kolmogorov-Arnold Networks 7364
Xiangfei Fang, Boying Wang, Chengying Huan, Shaonan Ma, Heng Zhang, Chen Zhao

Hybrid Contrastive Learning Decoupling Speech Emotion Recognition 7369
Chenyu Li, Yue Ma2 Gu, He Zhang, Linsong Liu, Haixiang Lin, Shuang Wang

Multi-Scale Dehaze Network Based on Frequency Domain Assistance and Detailed Brightness Information Guidance 7374
Pengwei Yang, Lei Wang

Zero-Shot Quantization for Large-Kernels Via Shape-Based Distribution and Diversity Self-Distillation 7379
Yao Li, Zhuozhen Yu, Xinrui Chen, Shunzhou Wang, Hang Yuan, Wei Gao

Steerable Differential Polynomial Beamforming 7384
Feng Jiang, Huawei Chen

Federated Hybrid-Supervised Learning for Universal Medical Image Segmentation 7389
Shenhai Zheng, Sian Wen, Congyu Li, Qing Chen, Laquan Li

Multi-Scale Attention-Based Dense Spatial-Temporal Model for Emotion Induction in Response to Olfactory Stimuli 7394
Jian-Ming Zhang, Wei-Bang Jiang, Wei-Long Zheng, Bao-Liang Lu

Diffusion Models Are Good Unsupervised Class-Agnostic Shape Part Segmentators 7399
Zhongbin Jiang, Tianhao Shi, Hao Gao, Jun Liu, Ye Liu

RefleXGen:The Unexamined Code is Not Worth Using 7404
Bin Wang, Hui Li, AoFan Liu, BoTao Yang, Ao Yang, YiLu Zhong, Weixiang Huang, Runhuai Huang, Weimin Zeng, Yanping Zhang

A Dual-Stream Network with Non-Stationary Characteristics-Enhanced for SST Image Prediction 7409
Min Ye, Jiahao Du, Ning Song, Xinyue Liang, Bohan Li, Jie Nie

Self-Supervised Contrastive Pre-Training for Dry Electrode EEG Emotion Recognition Via Cross Device Representation Consistency 7414
Meihong Zhang, Shaokai Zhao, Zhiguo Luo, Liang Xie, Tiejun Liu, Dezhong Yao, Ye Yan, Erwei Yin

Path Signatures Are Unsupervised Time Series Anomaly Extractors 7419
Ruiqi Wang, Zhenwei Zhang, Yuantao Gu

Wave-Spectrogram Cross-Modal Aggregation for Audio Deepfake Detection 7424
Zehui Jin, Linlong Lang, Biao Leng

LiDAR-SPD: Improving Adversarial Robustness of 3D Object Detection Via Spherical Projection and Diffusion 7429
Mumuxin Cai, Xupeng Wang, Ferdous Sohel, Hang Lei

LitePest: Real-Time and Efficient Detection of Agricultural Pests Using an Advanced Lightweight Deep Learning Network 7434
Zhe Tang, Jiajia Lu, Wei Xiang, Wanyu Ling, Lingyan Zhang

Improved Techniques for Offline Reinforcement Learning: Advantage Value Estimation and Layernorm 7439
Xiaosong Liu, Quan Liu, Lan Wu

Regularized Weighted Descent: Model-Based Learner for Multi-Target Radar Waveform Design 7444
Junho Kweon, Fulvio Gini, Maria S. Greco, Muralidhar Rangaswamy, Vishal Monga

Self-Supervised Image Harmonization Via Holistic Feature Fusion 7449
Chenyang Tian, Qing Zhang

Enhancing Federated Domain Adaptation Via Multi-Granular Fine-Grained Alignment 7454
Ziyun Cai, Shangshang Song, Jie Song, Yawen Huang, Changhui Hu, Xiao-Yuan Jing

DSDIR: A Two-Stage Method for Addressing Noisy Long-Tailed Problems in Malicious Traffic Detection 7459
Guoliang Li, Ruiqi Zhang, Zhe Sun, Lingkai Xing, Yu Zhang

Gaussian Difference: Find Any Change Instance in 3D Scenes 7464
Binbin Jiang, Rui Huang, Qingyi Zhao, Yuxiang Zhang

Enhancing Video-Text Matching Via Sparse Stratified Sampling 7469
Chenyang Lyu, Wenxi Li, Tianbo Ji, Liting Zhou, Pintu Lohar, Yi Yu, Longyue Wang

SelectiveFinetuning: Enhancing Transfer Learning in Sleep Staging Through Selective Domain Alignment 7474
Siyuan Zhao, Chenyu Liu, Yi Ding, Xinliang Zhou

Graph Learning with Low-Rank and Diagonal Structures: A Riemannian Geometric Approach 7479
Xiang Zhang, Qiao Wang

A Chinese Expressive Long-Dialogue Speech Dataset with Scripts 7484
Jin Li, Tianrui Wang, Meng Ge, Chenrui Cui, Jiahui Zhao, Jianrong Wang, Longbiao Wang, Jianwu Dang

PriorSinger: Singing Voice Synthesis Model with Prior Condition Cross Attention 7489
Zehua Zhang, Bosong Yan, Yinghan Cao, Mingjiang Wang

Semi-Supervised Iterative Learning Network for Camouflaged Object Detection 7494
Guowen Yue, Ge Jiao, Jiahao Xiang

SSR-Speech: Towards Stable, Safe and Robust Zero-Shot Text-Based Speech Editing and Synthesis 7499
Helin Wang, Meng Yu, Jiarui Hai, Chen Chen, Yuchen Hu, Rilin Chen, Najim Dehak, Dong Yu

DCFormer: Divide-And-Conquer in 3D Human Pose Estimation Tasks 7504
Tianyi Ma, Muqing Wu, Zijian Zhang

SoloAudio: Target Sound Extraction with Language-Oriented Audio Diffusion Transformer 7509
Helin Wang, Jiarui Hai, Yen-Ju Lu, Karan Thakkar, Mounya Elhilali, Najim Dehak

MACA: Multi-Anchor Classification Approach for Unsupervised Domain Adaptation 7514
Dexuan Zhao, Chong Zhao, Taizhang Hu, Xing Wei, Fan Yang, Yang Lu

ACRL-10K: A Dataset for Air Conditioner Refrigerant Leak Smoke Detection 7519
Yuhan Jiang, Zhongyuan Wang, Jie Hua, Jinbi Liang, Zengmin Xu

Radar2ECG: Multi-Scale Bottleneck Fusion and Cross-Modal Semantic Distillation for Conditional Electrocardiogram Generation from Radar Heart Sound 7524
Jinye Li, Aidong Men, Yang Liu, Pengda Han, Qingchao Chen

Macaque-Motion-Monitor Dataset: A New Benchmark for Macaque Action Recognition 7529
Jiawei Huang, Zhiyuan Chen, Wenxuan Fan, Xiaomei Zhang, Xibo Ma

Generalized Graph Signal Reconstruction Via the Uncertainty Principle 7534
Yanan Zhao, Xingchao Jian, Feng Ji, Wee Peng Tay, Antonio Ortega

RETAIN: Reliable Topology Augmentation for Both Heterophilic and Homophilic Graphs 7539
Ziyun Zou, Lian Shen, Yanhao Li, Yifan Lu, Juan Liu, Xiangrong Liu

MAP: Supporting Multimodal Knowledge Graph Completion Via Augmented Modality Alignment and Instance Preserving 7544
Yi Li, Qingmeng Zhu, Fei Song, Changwen Zheng, Jiangmeng Li

Bridging the Modality Gap for Speech-Image Retrieval with Text Supervision 7549
Yuting Yang, Lifeng Zhou, Yuke Li, Guodong Ma

Frequency-Domain Guided Multiple Parallel Kernels Network for Low-Light Remote Sensing Image Enhancement 7554
Jingxuan Zhou, Hao Li, Jinlong Wang, Xiongxin Tang, Fanjiang Xu

Classification of Eye-Tracking Data Based on Spatiotemporal Attention Encoding 7559
Jiaju He, Chen Xia, Kuan Li, Tian Zhang

LLDB: Efficient Low-Light Image Enhancement with Difffusion Bridge 7564
Junlong Ma, Conghan Yue, Zhengwei Peng, Dongyu Zhang

ITW-DehazeFormer: Imaging Through Turbid Water Using Improved DehazeFormer 7569
Qicong Wang, Xiaopin Zhong, Dajiang Lu, Yibin Tian

Towards Differential Optimization: Rehearsal-Free Class-Incremental Learning with Slow Learners and Fast Adapters 7574
Yinghong Chen, Huanjia Zhu, Jieyi Cai, Huanyu Liu, Jun Liang, Bingzhi Chen

Instance Segmentation of Airway Anatomies Using Mask R-CNN Prompt Adaptation-SAM 7579
Yinzhou Ling, Jingjing Luo, Yuan Han, Wenxian Li, Hongbo Wang

Semi-Supervised Speaker Diarization Using Graph Transformers and LLMs on Naturalistic Apollo 11 Data 7584
Meena M. C. Shekar, John H. L. Hansen

LOFI: Harnessing Attention Dynamics for Facial Expression Recognition with Noisy Labels 7589
Jinglin Zhang, Qiangchang Wang, Xinxin Zhang, Yilong Yin

Self-Support Prototype-Aware for Few-Shot Semantic Segmentation 7594
Jiaxiang Fang, Shiqiang Ma, Shengfeng He, Fei Guo

SAR Ship Detector Using Cross-Stage Feature Fusion and Decoupled Head with Mutual Guidance 7599
Yixin Qiao, Xiaoxiao Yin, Xinyuan Zhou, Shiyong Lan, Guonan Deng

Memory-Free and Parallel Computation for Quantized Spiking Neural Networks 7604
Dehao Zhang, Shuai Wang, Yichen Xiao, Wenjie Wei, Yimeng Shan, Malu Zhang, Yang Yang

Integrated Interpolation and Matrix Completion for Radio Map Estimation: A Convex Optimization Approach 7609
Hongcheng Dong, Wenqiang Pu, Rui Zhou, Xiao Fu, Feng Yin

IDE: A Multi-Agent-Driven Iterative Framework for Dynamic Evaluation of LLMs 7614
Xin Tong, Bo Jin, Jingya Wang, Wenpeng Xing, Tian Xia, Meng Han

Joint Beamforming Design for Multi-Functional RIS-Aided Over-The-Air Computation 7619
Xinran Zhang, Hui Tian, Wanli Ni

SCI-Gaussian: Optimizing 3D Gaussian Radiance Fields from a Snapshot Compressive Image 7624
Xiaoyue Li, Yunhao Li, Xiaodong Wang, Xin Yuan, Mark D. Butala, Gaoang Wang

ToolFiVe: Enhancing Tool-Augmented LLMs Via Tool Filtering and Verification 7629
Hailun Lu, Xingming Li, Xuanyu Ji, Zhigang Kan, Qingyong Hu

Robust Multi-Task Adversarial Attacks Using Min-Max Optimization 7634
Jiacheng Guo, Lei Li, Haochen Yang, Baocheng Geng, Hongkai Yu, Minghai Qin, Tianyun Zhang

Source-Free Domain Adaptation with Multiple Alignment for Efficient Image Retrieval 7639
Jinghui Ni, Hui Cui, Lihai Zhao, Fengling Li, Xiaohui Han, Lijuan Xu

Synthesizing Efficient Trajectory-Controllable Co-Speech Gesture with Latent Consistency Model 7644
Wenlong Wang, Dahua Gao, Yuxi Hu

Embedding Enhanced MLP Enables Simple and Extensible Spatiotemporal Forecasting 7649
Ziyan Wang, Yudong Zhang, Pengkun Wang, Xu Wang, Yang Wang

Improving Speech Enhancement by Cross- And Sub-Band Processing with State Space Model 7654
Jizhen Li, Weiping Tu, Yuhong Yang, Xinmeng Xu, Yiqun Zhang, Yanzhen Ren

Joint Automatic Speech Recognition and Structure Learning for Better Speech Understanding 7659
Jiliang Hu, Zuchao Li, Mengjia Shen, Haojun Ai, Sheng Li, Jun Zhang

CLAP-S: Support Set Based Adaptation for Downstream Fiber-Optic Acoustic Recognition 7664
Jingchen Sun, Shaobo Han, Wataru Kohno, Changyou Chen

HATTM: A Novel Hybrid Attention Model for Ethereum Phishing Scams Detection 7669
Bo Cui, Zhenyu Zhang, Wenhan Hou

Integrated Global-Local Gaussian Attention for Image Compression 7674
Atefeh Khoshkhahtinat, Piyush M. Mehta

Similarity-Based Accent Recognition with Continuous and Discrete Self-Supervised Speech Representations 7679
Jun-You Wang, Sheng Li, Li-An Lu, Sydney Chia-Chun Kao, Jyh-Shing Roger Jang

Multi-Modal Semantic Communication with Point Cloud Diffusion 7684
Weiyan Feng, Zhenyu Shi

Fine-Grained Vital Sign Reconstruction Through Machine Learning on Multi-Channel Radar Signals 7689
Changming Li, Cong Shi, Athina Petropulu, Yingying Chen

KARLM: Enhancing LLM-Based Recommendation Systems with Knowledge Bases 7694
Ze Song, Dehong Chen, Xiaoyi Shen, Xiangyu Zhou, Ji Qi, Yi Zhou

DDSP Guitar Amp: Interpretable Guitar Amplifier Modeling 7699
Yen-Tung Yeh, Yu-Hua Chen, Yuan-Chiao Cheng, Jui-Te Wu, Jun-Jie Fu, Yi-Fan Yeh, Yi-Hsuan Yang

Exploiting Attention-To-Motion Via Transformer for Versatile Video Frame Interpolation 7704
Chee-Kim Gan, Jian-Jiun Ding, Chang-Yu Hsieh, De-Yan Lu

Audiogram-Informed End-To-End Noise Reduction and Wide Dynamic Range Compression for Hearing Aids 7709
Huiyong Zhang, Brian C. J. Moore, Lingling Dai, Fengyuan Hao, Xiaodong Li, Chengshi Zheng

Cross-Template-Based Hypergraph Transformer 7714
Jingxi Feng, Xiangmin Han, Heming Xu, Juan Wang, Jue Jiang, Shaoyi Du, Yue Gao

ViM-Disparity: Bridging the Gap of Speed, Accuracy and Memory for Disparity Map Generation 7719
Maheswar Bora, Tushar Anand, Saurabh Atreya, Aritra Mukherjee, Abhijit Das

Speaking from Coarse to Fine: Improving Neural Codec Language Model Via Multi-Scale Speech Coding and Generation 7724
Haohan Guo, Fenglong Xie, Dongchao Yang, Xixin Wu, Helen Meng

Exploring Graph-Aware Reasoning and Bidirectional Selection for Vision-Language Navigation 7729
Dongming Zhou, Jinsheng Deng, Zhengbin Pang, Wei Li

KVPruner: Structural Pruning for Faster and Memory-Efficient Large Language Models 7734
Bo Lv, Quan Zhou, Xuanang Ding, Yan Wang, Zeming Ma

Decision-Aided Progressive Symbol Phase Equalizer in Sweep Spread Carrier Underwater Acoustic Communications 7739
Anoop R., Manju M. Raj, Arunkumar K. P., Chandra R. Murthy

Formula-Supervised Sound Event Detection: Pre-Training Without Real Data 7744
Yuto Shibata, Keitaro Tanaka, Yoshiaki Bando, Keisuke Imoto, Hirokatsu Kalaoka, Yoshimitsu Aoki

Identical Human Preference Alignment Paradigm for Text-To-Image Models 7749
Haoyuan Sun, Bo Xia, Yifei Zhao, Yongzhe Chang, Xueqian Wang

Col-OLHTR: A Novel Framework for Multimodal Online Handwritten Text Recognition 7754
Chenyu Liu, Jinshui Hu, Baocai Yin, Jia Pan, Bing Yin, Jun Du, Qingfeng Liu

KLMN: Knowledge Distillation Based Lightweight Multi-Clue Image Forgery Detection and Localization 7759
Heng Huang, Yaqi Liu, Xin Jin, Song Xiao, Bin Liu

TheSHY-3D: Texture and Structure HarmonY for Multi-View 3D Object Detection 7764
Wenquan Zhang

An Ensemble Approach to Short-Form Video Quality Assessment Using Multimodal LLM 7769
Wen Wen, Yilin Wang, Neil Birkbeck, Balu Adsumilli

MDN: Mamba-Driven Dualstream Network for Medical Hyperspectral Image Segmentation 7774
Shijie Lin, Boxiang Yun, Wei Shen, Qingli Li, Anqiang Yang, Yan Wang

FruitMMBench: A Multi-Modal Benchmark for Fruit Quality Assessment 7779
Jiawei Chen, Gong Huang, Liu Liu, Zhenbo Xu, Qinghong Yang

Multichannel-To-Multichannel Target Sound Extraction Using Direction and Timestamp Clues 7784
Dayun Choi, Jung-Woo Choi

Diffusion-Based Target Device Style Transfer for Robust Acoustic Scene Classification 7789
Won-Gook Choi, Joon-Hyuk Chang

Hybrid Coding and Weakly-Supervised Approach for Depth Estimation from Wrapped Phase 7794
Jie Ren, Chunqian Tan, Wanzhong Song

Self-Supervised Learning for Acoustic Few-Shot Classification 7799
Jingyong Liang, Bernd Meyer, Issac Ning Lee, Thanh-Toan Do

PBD : Plastic Bottle Dataset for Defect Detection 7804
Zhenyuan Lin, Danhua Liu, Lai Wei, Yubo Dong

Mora-Level Prosody Prediction for Text-To-Speech Using Japanese BERT Without Accentual Labels 7809
Tadashi Ogura, Takuma Okamoto, Yamato Ohtani, Erica Cooper, Tomoki Toda, Hisashi Kawai

Predicting Local fMRI Activations from EEG: A Feasibility Study Using Both Classical and Modern Machine Learning Pipelines 7814
Tomer Amit, Taly Markovits, Guy Gurevitch, Talma Hendler, Lior Wolf

CMoS: Customizing Model Structures for Personalized Federated Learning 7819
Bingyan Liu, Jing Yang

Boosting Movie and TV Tag Accuracy with Knowledge Graphs 7824
Hongxun Jiang, Lin Zhang, Lifeng Zhang

Unleashing Potential of Evidence in Knowledge-Intensive Dialogue Generation 7829
Xianjie Wu, Jian Yang, Tongliang Li, Shiwei Zhang, Yiyang Du, LinZheng Chai, Di Liang, Zhoujun Li

Longitudinal Wrist PPG Analysis for Reliable Hypertension Risk Screening Using Deep Learning 7834
Hui Lin, Jiyang Li, Ramy Hussein, Xin Sui, Xiaoyu Li, Guangpu Zhu, Aggelos K. Katsaggelos, Zijing Zeng, Yelei Li

Transformer-Based Contrastive Meta-Learning for Low-Resource Generalizable Activity Recognition 7839
Junyao Wang, Mohammad Abdullah Al Faruque

Stable Reduced-Rank VAR Estimators in Closed Forms 7844
Xinhui Rong

YOLO-TCT: An Effective Network for Long-Tailed Cervical Cell Detection 7849
Di Lv, Lin Yi, Li Liu, Yuze Chen, Xin Chen, Ran Liu

FUVAS: Few-Shot Unsupervised Video Anomaly Segmentation Via Low-Rank Factorization of Spatio-Temporal Features 7854
Jiaxiang Jiang, Ibrahima J. Ndiour, Mahesh Subedar, Omesh Tickoo

Matrix Decomposition by Additive and Subtractive Factors 7859
Takumi Kobayashi, Kenji Watanabe

Full-Reference Point Cloud Quality Assessment with Multimodal Large Language Models 7864
Ryosuke Watanabe, Tomoaki Konno, Hiroshi Sankoh, Bryan Tanaka, Tatsuya Kobayashi

Deep Support Vein Machine for Lung Parcellation 7869
Haichao Peng, Hao Fang, Wenkang Fan, Yong Wang, Sunkui Ke, Jie Luo, Xiongbiao Luo

Reversible Data Hiding in Encrypted Images Based on Variational Lossless Compression and Dual Encryption 7874
Yaolin Yang, Hongjie He

Distance Based Single-Channel Target Speech Extraction 7879
Runwu Shi, Benjamin Yen, Kazuhiro Nakadai

HyperSMOTE: A Hypergraph-Based Oversampling Approach for Imbalanced Node Classifications 7884
Ziming Zhao, Tiehua Zhang, Zijian Yi, Zhishu Shen

Representative Arm Identification: A Fixed Confidence Approach to Identify Cluster Representatives 7889
Sarvesh Gharat, Aniket Yadav, Nikhil Karamchandani, Jayakrishnan Nair

Incorporate Global Information from Entire Datasets for Knowledge Tracing Via Mini-Batch Input 7894
Hui Zhao, Tingyu Fu

AuscMLLM: Bridging Classification and Reasoning in Heart Sound Analysis with a Multimodal Large Language Model 7899
Zihan Zhao, Pingjie Wang, Liudan Zhao, Ya Zhang, Kun Sun, Xin Sun, Xin Zhou, Yanfeng Wang, Yu Wang

Perception-Enhanced Network for Accurate Human Pose Estimation 7904
Xiaodi Sun, Baojiang Zhong, Kai-Kuang Ma

QoKV: Comprehending and Surpassing the Hurdles of KV Cache Quantization 7909
Jinguang Wang, Yuexi Yin, Haifeng Sun, Tingting Yang, Qi Qi, Zirui Zhuang, Jingyu Wang

Training-Free Point Cloud Recognition Based on Geometric and Semantic Information Fusion 7914
Yan Chen, Di Huang, Zhichao Liao, Xi Cheng, Xinghui Li, Long Zeng

Less is More: Embracing Sparsity and Interpolation with Esiformer for Time Series Forecasting 7919
Yangyang Guo, Yanjun Zhao, Sizhe Dang, Tian Zhou, Liang Sun, Yi Qian

Learning Hierarchical Attribute Prompt for Vision-Language Models 7924
Jun Liang, Yang Peng, Rui Luo, Yunyu Zou, Yalong Cheng, Bingzhi Chen

Multi-Layer Knowledge Distillation for Continual Semantic Segmentation 7929
Yan Wang, Pengju Xu, Bingye Wang, Haiying Zhao

Heterogeneous Data-Based Cross-Domain Few-Shot Classification Method of Hyperspectral Image 7934
Kun Zhao, Shaoguang Huang, Hongyu Chen, Hongyan Zhang

Audio Texture Manipulation by Exemplar-Based Analogy 7939
Kan Jen Cheng, Tingle Li, Gopala Anumanchipalli

Enhancing Autonomous Driving Through Dual-Process Learning with Behavior and Reflection Integration 7944
Xiao Zhang, Kangsheng Wang, Tianyu Hu, Huimin Ma

MAJoR: Visual Emotion Analysis Via Multi-Attribute Joint Reasoning 7949
Yuxin Fei, Jinlan Xu, Maoying Qiao, Fei Gao

DFMA: Adaptive Dual Fusion for Multimodal Relation Extraction with Mutual Attention 7954
Yunhao Li, Tianxi Huang, Guiduo Duan, Yunrui Dong

Enhancing Autonomous Vehicle Planning with a Robust Fault-Tolerant Mechanism for Action-Induced Agent Detection 7959
Zheng Fu, Hezhe Lin, Kangan Qian, Tuopu Wen, Hao Gao, Zhihua Zhong, Diange Yang

DuCol: Text-Tag Adaptive Colorization of Dual-Character Line Art 7964
Jun Liang, Rui Luo, Yang Peng, Hai Su

Convolutional Retentive Network for EEG Decoding 7969
Junliang Wang, Wenlong Hang, Shuang Liang, Qiong Wang, Badong Chen, Jing Qin

HFedPFS: Heterogeneous Federated Learning with Personalized Data Feature Sharing 7974
Jingxian Xu, Liping Yi, Gang Wang, Xiaoguang Liu

Cross-Layer Graph Knowledge Distillation for Image Recognition 7979
Jiaming Chu, Yanzhuo Xiang, Yuqi Li, Chuanguang Yang, Zhulin An, Yongjun Xu

Voice Conversion for Low-Resource Languages Via Knowledge Transfer and Domain-Adversarial Training 7984
Huu Tuong Tu, Luong Thanh Long, Vu Huan, Nguyen Thi Phuong Thao, Nguyen Van Thang, Nguyen Tien Cuong, Nguyen Thi Thu Trang

FAST: Fast Audio Spectrogram Transformer 7989
Anugunj Naman, Gaibo Zhang

MIB: Mixed Information Bottleneck for Out-Of-Distribution Keyword Spotting 7994
Yashas Malur Saidutta, Rakshith Sharma Srinivasa, Jaejin Cho, Ching-Hua Lee, Chouchang Yang, Yilin Shen, Hongxia Jin

Projection Valued-Based Quantum Machine Learning Adapting to Differential Privacy Algorithm for Word-Level Lipreading 7999
Hang Chen, Chang Wang, Jun Du, Chao-Han Huck Yang, Jun Qi

Augmenting Short Enrollment Speech Via Synthesis for Target Speaker Extraction 8004
Zikang Huang, Jingru Lin, Meng Ge, Yu Jiang, Xiaobao Wang, Longbiao Wang, Jianwu Dang

Graph Signal Reconstruction Via Koopman Autoencoder 8009
Sivaram Krishnan, Jihong Park, Jinho Choi

Exploiting the Relationship Within the Unlabelled Samples by Set Matching for Generalized Category Discovery 8014
Qiubo Ma, Hang Yu, Yuan Shan, Pinzhuo Tian

FADEL: Uncertainty-Aware Fake Audio Detection with Evidential Deep Learning 8019
Ju Yeon Kang, Ji Won Yoon, Semin Kim, Min Hyun Han, Nam Soo Kim

SoundCollage: Automated Discovery of New Classes in Audio Datasets 8024
Ryuhaerang Choi, Soumyajit Chatterjee, Dimitris Spathis, Sung-Ju Lee, Fahim Kawsar, Mohammad Malekzadeh

Adaptive Feature Aggregation for In-Air Handwritten Trajectory 8029
Zeyu Qiu, Weiqiang Wang

TD-RD: A Top-Down Benchmark with Real-Time Framework for Road Damage Detection 8034
Xi Xiao, Zhengji Li, Wentao Wang, Jiacheng Xie, Houjie Lin, Swalpa Kumar Roy, Tianyang Wang, Min Xu

Target Word Activity Detector: An Approach to Obtain ASR Word Boundaries Without Lexicon 8039
Sunit Sivasankaran, Eric Sun, Jinyu Li, Yan Huang, Jing Pan

EventLens: Enhancing Visual Commonsense Reasoning by Leveraging Event-Aware Pretraining and Cross-Modal Linking 8044
Mingjie Ma, Zhihuan Yu, Yichao Ma, Guohui Li, Zhong Yang

EDSep: An Effective Diffusion-Based Method for Speech Source Separation 8049
Jinwei Dong, Xinsheng Wang, Qirong Mao

Chain-Of-Thought Prompting for Speech Translation 8054
Ke Hu, Zhehuai Chen, Chao-Han Huck Yang, Piotr Żelasko, Oleksii Hrinchuk, Vitaly Lavrukhin, Jagadeesh Balam, Boris Ginsburg

Exploring Acoustic Foundations in Speech Production Assessment Models for Children with Cochlear Implants 8059
Seonwoo Lee, Sunhee Kim, Minhwa Chung

Diffusion Models Are Zero-Shot Generative Text-Vision Retrievers 8064
Bao Li, Zeke Xie, Xiaomei Zhang, Xiangyu Zhu, Zhen Lei

Enhancing Multivariate Time Series Forecasting with Multi-Scale Moving Transformation 8069
Wenjie Ou, Hongmin Du, Wenqiang Zhao, Dongyue Guo, Yi Lin

Plaintext-Free Deep Learning for Privacy-Preserving Medical Image Analysis Through Frequency Information Embedding 8074
Mengyu Sun, Ziyuan Yang, Maosong Ran, Zhiwen Wang, Hui Yu, Yi Zhang

Mask-Guided Multi-Scale Spatial-Spectral Transformer for Snapshot Compressive Imaging 8079
Heyuan Yin, Jingxiang Yang, Jia Liu, Liang Xiao

Towards a Distribution Alignment Framework for Incomplete Data Classification 8084
Linqing Huang, Jinfu Fan, Shilin Wang, Gongshen Liu, Shouxuan Liu

Attention-Enhanced Short-Time Wiener Solution for Acoustic Echo Cancellation 8089
Fei Zhao, Xueliang Zhang

CAMEL: Cross-Attention Enhanced Mixture-Of-Experts and Language Bias for Code-Switching Speech Recognition 8094
He Wang, Xucheng Wan, Naijun Zheng, Kai Liu, Huan Zhou, Guojian Li, Lei Xie

SelaFD:Seamless Adaptation of Vision Transformer Fine-Tuning for Radar-Based Human Activity Recognition 8099
Yijun Wang, Yong Wang, Chendong Xu, Shuai Yao, Qisong Wu

A Modified Gain Normalized Step Size Adaptive Algorithm for Improved Online Secondary Path Modelling in Active Noise Control 8104
Asutosh Kar, Gagandeep Singh, Pradeep K. Shill, Somanath Pradhan, Vasundhara, M. G. Christensen

Tribe Graph Enhanced Bidirectional Mamba for Multivariate Time Series Forecasting 8109
Yingqi Zhao, Haiwei Zhang, Jiaqi Ye, Shubao Zhao, Chengyi Yang, Zengxiang Li

KCE-Unet: A Novel Music Denoising Method with KANConv ECA Unet* 8114
Shijie Zhang, Yulun Wu, Ganghui Ru, Yi Yu, Wei Li

RAW Data: A Key Component for Effective Deepfake Detection 8119
Sahar Husseinil, Jean-Luc Dugelay

MHSDB: A Comprehensive Benchmark for Multimodal Humor and Sarcasm Detection Leveraging Foundation Models..... 8124
Zhongren Dong, Donghao Wang, Ciqiang Chen, Dong-Yan Huang, Zixing Zhang

SEAL: Speaker Error Correction Using Acoustic-Conditioned Large Language Models..... 8129
Anurag Kumar, Rohit Paturi, Amber Afshan, Sundararajan Srinivasan

Retrieval-Augmented Neural Field for HRTF Upsampling and Personalization 8134
Yoshiki Masuyama, Gordon Wichern, François G. Germain, Christopher Ick, Jonathan Le Roux

Automatic Speech Recognition and Spoken Language Understanding of Maritime Radio Communications: A Case Study with Singapore Data..... 8139
Phuong Dat, Jayakrishnan Melur Madhathil, Tran Huy Dat

The CDC Problem: Distributed Spatial Sampling and Detection of Poisson Processes..... 8144
Vanlalruata Ralte, Amitalok J. Budkuley, Stefano Rini

MRI2Speech: Speech Synthesis from Articulatory Movements Recorded by Real-Time MRI 8149
Neil Shah, Ayan Kashyap, Shirish Karande, Vineet Gandhi

Position-Aware Hypergraph Message-Passing Neural Network 8154
Xinyu Zhang, Qize Jiang, Hanyuan Zhang, Weiwei Sun

MmHIU: A Human-To-Human Interaction Understanding System Based on mmWave Sensing 8159
Fenglin Zhang, Chenglin Wu, Anfu Zhou, Huadong Ma

Generalized Approximate Message-Passing for Compressed Sensing with Sublinear Sparsity..... 8164
Keigo Takeuchi

SegAug: CTC-Aligned Segmented Augmentation for Robust RNN-Transducer Based Speech Recognition 8169
Khanh Le, Tuan Vu Ho, Dung Tran, Duc Thanh Chau

LlamaPartialSpoof: An LLM-Driven Fake Speech Dataset Simulating Disinformation Generation 8174
Hieu-Thi Luong, Haoyang Li, Lin Zhang, Kong Aik Lee, Eng Siong Chng

Basket-Enhanced Heterogenous Hypergraph for Price-Sensitive Next Basket Recommendation 8179
Yuening Zhou, Yulin Wang, Qian Cui, Xinyu Guan, Francisco Cisternas

Generate E-Commerce Product Background by Integrating Category Commonality and Personalized Style 8184
Haohan Wang, Wei Feng, Yaoyu Li, Zheng Zhang, Jingjing Lv, Junjie Shen, Zhangang Lin, Jingping Shao

MILE: Multi-Instance Learning for Document Event Argument Extraction 8189
Jiaxian Wang, Yong Zhang, Xiang Peng

UML: A Unified Multimodal Learning Framework for Cataract Postoperative Visual Acuity Prediction with Uncertain Missing Modalities 8194
Tongyu Yang, Qian Zhou, Hua Zou, Haifeng Jiang, Yong Wang

Exploring Large Language Models for Knowledge Graph Completion 8199
Liang Yao, Jiazhen Peng, Chengsheng Mao, Yuan Luo

SPT: Sequence Prompt Transformer for Interactive Image Segmentation 8204
Senlin Cheng, Haopeng Sun, Tao Xie, Hangyue Zhao, Yiqiang Chen, Bolei Xu, XiaoBo Li

ChunkFormer: Masked Chunking Conformer for Long-Form Speech Transcription 8209
Khanh Le, Tuan Vu Ho, Dung Tran, Duc Thanh Chau

RADCI: A Synchronized Radar-RGBT Object Detecting-Tracking Dataset and a Benchmark 8214
Heng Yu, Ruiheng Zhang, Haoyang Sun, Zhe Cao, Biwen Yang, Jin Zhang, Guanyu Liu

What Affects the Performance of Fake Audio Detection? Analyzing Factors in a Continual Learning Setting 8219
Yixuan Xiao, Ngoc Thang Vu

Low Complexity DoA-ToA Signature Estimation for Multi-Antenna Multi-Carrier Systems 8224
Chandrashekhar Rai, Debarati Sen

Bayesian Nonparametric Clustering for Source Counting with a Small Aperture Microphone Array 8229
Kunkun SongGong, Pufen Zhang, Xiongwei Zhang, Wenwu Wang, Meng Sun, Chong Jia, Yihao Li

O-EENC-SD: Efficient Online End-To-End Neural Clustering for Speaker Diarization 8234
Elio Gruttadauria, Mathieu Fontaine, Jonathan Le Roux, Slim Essid

Rethinking Mean Opinion Scores in Speech Quality Assessment: Score Aggregation Through Quantized Distribution Fitting 8239
Yuto Kondo, Hirokazu Kameoka, Kou Tanaka, Takuhiro Kaneko

KSSANet: KAN-Driven Spatial-Spectral Attention Networks for Hyperspectral Image Super-Resolution 8244
Baisong Li, Xingwang Wang, Haixiao Xu

Object-Based Video Tampering Localization Via Trace Consistency Analysis 8249
Pengfei Pei, Yun Cao, Jinchuan Li, Zeyu Zhang, Yuqi Pang

SFFCE-CD: Spatial and Frequency Feature Cross Enhancement for Change Detection 8254
Yan Xing, Jiali Hu, Binbin Jiang, Qingyi Zhao, Longxi Feng, Rui Huang

Topology Decoupled All-Reduce Algorithm 8259
Ruixing Zong, Jiapeng Zhang, Zhuo Tang, Anwitaman Datta

PrimeK-Net: Multi-Scale Spectral Learning Via Group Prime-Kernel Convolutional Neural Networks for Single Channel Speech Enhancement 8264
Zizhen Lin, Junyu Wang, Ruili Li, Fei Shen, Xi Xuan

KNN-SVC: Robust Zero-Shot Singing Voice Conversion with Additive Synthesis and Concatenation Smoothness Optimization 8269
Keren Shao, Ke Chen, Matthew Baas, Shlomo Dubnov

PR-KGC: Text-Enhanced Knowledge Graph Completion with Pair-Wise Re-Ranking 8274
Meng Gao, Feng Zhang, Wei Chen, Tengjiao Wang, Jiabin Zheng, Jiahui Yao

GPA: Enhancing Generalizable Physical Adversarial Attacks Across Multiple Vision Tasks 8279
Mingye Xie, Suncheng Xiang, Jiacheng Ruan, Xian Gao, Zefang Yu, Ting Liu, Yuzhuo Fu

LCFed: An Efficient Clustered Federated Learning Framework for Heterogeneous Data 8284
Yuxin Zhang, Haoyu Chen, Zheng Lin, Zhe Chen, Jin Zhao

Robust Deterministic DOA Estimation Using α-Divergence in Unknown Noise Fields with Sparse Sensor Arrays 8289
Wenjing Yang, Abd-Krim Seghouane, Pavel Krupskiy

An Optimized GPU-Based Acceleration of CRYSTALS-Dilithium 8294
Zhao Chen, Weimin He, Jiafei Wu, Jingjie Liu, Boqin Xu, Xiaoning Bian, Zhe Liu

The Sound of Language: A Bilingual Analysis of Voice Conversion and Text-To-Speech Synthesis 8299
Jeong-Eun Choi, Karla Schäfer, Martin Steinebach

StableQuant: Layer Adaptive Post-Training Quantization for Speech Foundation Models 8304
Yeona Hong, Hyewon Han, Woo-Jin Chung, Hong-Goo Kang

MFDPonzi: Detecting Ethereum Ponzi Schemes Using Static Features from Novel Opcode Sequences 8309
Longwei Cao, Jiwei Qin, Xuzi Zhang

Diff-SAGe: End-To-End Spatial Audio Generation Using Diffusion Models 8314
Saksham Singh Kushwaha, Jianbo Ma, Mark R. P. Thomas, Yapeng Tian, Avery Bruni

Mamba-SEUNet: Mamba UNet for Monaural Speech Enhancement 8319
Junyu Wang, Zizhen Lin, Tianrui Wang, Meng Ge, Longbiao Wang, Jianwu Dang

GENIE: Socially Unbiased Generative Text-To-Image Editing 8324
Julia K. Lau, Raphaël C.-W. Phan, Sailaja Rajanala, Ingemar J. Cox, Arghya Pal

Gradient-Oriented Clustered Federated Learning with Efficient Knowledge Sharing in Non-IID Settings 8329
Kenta Kubota, Ren Togo, Keisuke Maeda, Takahiro Ogawa, Miki Haseyama

Accompaniment Prompt Adherence: A Measure for Evaluating Music Accompaniment Systems 8334
Maarten Grachten, Javier Nistal

Retention Enhanced Cross-Modal Attention for Multi-Hop VQA 8339
Zijie Zhu, Feng Ding, Chenglong Chu, Fangming Zhong

ES-NeRF: Enhancing Segmentation in NeRF with CLIP 8344
Chong Zhao, Pengcheng Hou, Yan Zhai, Xing Wei, Fan Yang, Xiang Bi

ClingTP: Curriculum Learning Based Multi-Style Title Prefix Generation 8349
Yusong Wang, Dongyuan Li, Jialun Shen, Yicheng Xu, Shuai Zhong, Mingkun Xu

Cooperative and Competitive Functional Connectivity Based on Improved Ising Model 8354
Gengqian Wei, Chuang Liang, Tülay Adali, Rongtao Jiang, Daoqiang Zhang, Vince D. Calhoun, Shile Qi

A Singing Melody Extraction Network Via Self-Distillation and Multi-Level Supervision 8359
Ying Hu, Jiabo Jing, Fan Li, Lijun He, Li Lin, Wenzhong Yang

Critically-Damped Third-Order Langevin Dynamics 8364
Benjamin Sterling, Mónica F. Bugallo

Integrating Adaptive Sampling for Optimal Learned Video Compression 8369
Wuyang Cong, Yuzhuo Kong, Ming Lu, Lizhong Wang, Weijing Shi, Zhan Ma

Triplet Synthesis for Enhancing Composed Image Retrieval Via Counterfactual Image Generation 8374
Kenta Uesugi, Naoki Saito, Keisuke Maeda, Takahiro Ogawa, Miki Haseyama

LGNet: Linear Graph Representation for Efficient Cold-Start Recommendations 8379
Zhenglong Li, Ruiqi Luo, Bangchao Wang, Lin Li, Xian Zhong

Multimodal Atrial Fibrillation Risk Stratification: Fusing Post-Stroke Brain DWI and Clinical Data 8384
Mohammad Javad Shokri, Nandakishor Desai, Aravinda S. Rao, Angelos Sharobeam, Bernard Yan, Marimuthu Palaniswami

Explicit Mutual Information Maximization for Self-Supervised Learning 8389
Lele Chang, Peilin Liu, Qinghai Guo, Fei Wen

Point Cloud Registration Via Reconstruction with Local Geometry Information Aggregation 8394
Zewei Pan, Linsen Li, Tongxin Yuan, Jiarong Yang

MQVAE: Capturing Metastable Dynamics from EEG for Brain-Computer Interfaces 8399
Yihang He, Chenfei Ye, Guoqing Cai, Jiahui Liao, Ting Ma

SentiFormer: Metadata Enhanced Transformer for Image Sentiment Analysis 8404
Bin Feng, Shulan Ruan, Mingzheng Yang, Dongxuan Han, Huijie Liu, Kai Zhang, Qi Liu

Curriculum Learning Aided Audio-Visual Speech Recognition with Arbitrary Speaker Number 8409
Yuxiao Lin, Tao Jin, Xize Cheng, Zhou Zhao, Fei Wu

Neuromorphic Unlimited Sampling for High-Dynamic-Range Video Acquisition 8414
Abijith Jagannath Kamath, Chandra Sekhar Seelamantula

Domain-Independent Automatic Generation of Descriptive Texts for Time-Series Data 8419
Kota Dohi, Aoi Ito, Harsh Purohit, Tomoya Nishida, Takashi Endo, Yohei Kawaguchi

Co-Attention Based Multi-Channel TF-GridNet for Speech Separation with Ad-Hoc Microphone Arrays 8424
Hongmei Guo, Linfeng Feng, Yijiang Chen, Xueqing Li, Boyu Zhu, Hao-Yu Wang, Xiao-Lei Zhang, Xuelong Li

Homogeneous Graph Extraction: An Approach to Learning Heterogeneous Graph Embedding 8429
Shihao Gao, Xiaoyan Yu, Yu Cai, Xulong Zhang, Jianzong Wang, Taisong Jin

Diversity Seeking Techniques for Red-Teaming Large Language Models 8434
Seokhan Lee, Bonhwa Ku, Hanseok Ko

Multiview Canonical Correlation Analysis for Automatic Pathological Speech Detection 8439
Yacouba Kaloga, Shakeel A. Sheikh, Ina Kodrasi

Cross-Modality Fusion Mamba for All-In-One Extreme Weather-Degraded Image Restoration 8444
Jiangang Ding, Yihui Shan, Lili Pei, Yiquan Du, Yuanlin Zhao, Wei Li

Non-Pharmacological Interventions: A Virtual Training Framework for Fine Motor Learning 8449
Xiaohang Dong, Qicheng Li, Yawen Z

Sociologically-Informed Graph Neural Network for Opinion Prediction 8454
Fan Yang, Jie Bai, Linjing Li, Daniel Zeng

Dynamically Optimize MTD Strategy in Satellite Computing Systems Using A2C Reinforcement Learning 8459
Lin Zhang, Yunchuan Guo, Shoukun Guo, Fenghua Li, Faqun Jiang, Liang Fang

PASTD: Progressive Augmentation and Spatiotemporal Decoupling Contrastive Learning for Skeleton-Based Action Recognition 8464
Qian Huang, Weiwen Qian, Chang Li, Gongyou Xu, Zhongqi Chen

Unsupervised Learning for Gain-Phase Impairment Calibration in ISAC Systems 8469
José Miguel Mateos-Ramos, Christian Häger, Musa Furkan Keskin, Luc Le Magoarou, Henk Wymeersch

LoRATEE: A Secure and Efficient Inference Framework for Multi-Tenant LoRA LLMs Based on TEE 8474
Zechao Lin, Sisi Zhang, Xingbin Wang, Yulan Su, Yan Wang, Rui Hou, Dan Meng

Enhancing Expressive Voice Conversion with Discrete Pitch-Conditioned Flow Matching Model 8479
Jialong Zuo, Shengpeng Ji, Minghui Fang, Ziyue Jiang, Xize Cheng, Qian Yang, Wenrui Liu, Guangyan Zhang, Zehai Tu, Yiwen Guo, Zhou Zhao

A Hierarchical Reasoning Framework for Complex Question Answering Over Knowledge Graph with Reinforcement Learning 8484
Zhiqiang Zhang, Zhiyi Zhang, Yunxiao Zhang, Wen Zhao

An Efficient Pore Annotation Framework for Tight Sandstone Images with Segment Anything Model 8489
Dongsheng Li, Chunyan Zang, Huijie Zhang, Yiming Lin, Qiushi Xia

Personalized Speech Enhancement Without User Enrollment for Real-World Audio Replay Scenarios 8494
Haoran Wei, Shilin Wang, Yanhua Long

MMTP: Meta-Learning-Based Multi-Textual Prompt Tuning for Visual-Language Models 8499
Fangtong Sun, Junjie Zhu, Zunlin Fan, Yiying Li, Zhiyuan Wang, Ke Yang

TDMF: Text-Guided Denoising and Interactive Medical Image Fusion 8504
Aimei Dong, Jingyuan Xu, Long Wang, Guohua Lv, Guixin Zhao, Jinyong Cheng

DeFT-Mamba: Universal Multichannel Sound Separation and Polyphonic Audio Classification 8509
Dongheon Lee, Jung-Woo Choi

CSMT: Combining Snoring and Metadata-Based Text for Sleep Apnea Severity Classification 8514
Heng Li, Yukun Qian, Yun Lu, Mingjiang Wang

Non-Negative Weighted DAG Structure Learning 8519
Samuel Rey, Seyed Saman Saboksayr, Gonzalo Mateos

PulmoScan: A Practical Pulmonary Disease Pre-Screening System 8524
Baixu Yan, Shijia Ge, Meizi Lu, Weixiang Zhang, Shuzhao Xie, Zhi Wang

MoHGNN: Enhanced Heterogeneous Graph Neural Network Via Metapath Optimization 8529
Taiyao Zhang, Xingyu Fu, Yuxin Zhang, Qingyun Liu

Cloth-Debiasing with Stable Diffusion in Cloth-Changing Person Re-Identification 8534
Haiyang Zhang, Xinshuang Wang

A Plug-And-Play Diffusion-Styled Conversion Model for Domain Discrepancies in Medical Image Segmentation 8539
Dong Liu, Zhiyong Wang, Linlin Guo

Rhythmic Foley: A Framework for Seamless Audio-Visual Alignment in Video-To-Audio Synthesis 8544
Zhiqi Huang, Dan Luo, Jun Wang, Huan Liao, Zhiheng Li, Zhiyong Wu

MLSwinTNet: A Multi-Level Feature Interaction Network for Low-Light Image Enhancement 8549
Mingyang Sun, Xinxin Wang, Ru Yi

Fast and High-Quality Auto-Regressive Speech Synthesis Via Speculative Decoding 8554
Bohan Li, Hankun Wang, Situo Zhang, Yiwei Guo, Kai Yu

Robust Adversarial Defense Based on Non-Transferability of Attack Across Foundation Models 8559
Koshiro Toishi, Keisuke Maeda, Ren Togo, Takahiro Ogawa, Miki Haseyama

Reference-Guided Parallel Independent Component Analysis: Estimating Cognition Associated Multimodal Patterns in Schizophrenia 8564
Jingxian Hu, Chuang Liang, Tülay Adali, Qi Zhu, Daoqiang Zhang, Rongtao Jiang, Vince D. Calhoun, Shile Qi

Unbalanced Co-Relational Optimal Transport for Robust Heterogeneous Data Alignment 8569
Fengjiao Gong, Zichong Wang, Hongteng Xu

Metadata-Enhanced Speech Emotion Recognition: Augmented Residual Integration and Co-Attention in Two-Stage Fine-Tuning 8574
Zixiang Wan, Ziyue Qiu, Yiyang Liu, Wei-Qiang Zhang

Radar Jamming Recognition Via Cross-Modality Contrast Learning 8579
Ganggang Dong, Zixuan Wang

Known-Plaintext Attacks to Thumbnail-Preservation Encryption Using Pix2pix Generative Adversarial Network 8584
Zhiyang Li, Dong Xie, Sanqiang Liu, Fulong Chen, Peng Hu, Taochun Wang

Robust Exponential Hyperbolic Tangent Geman-McClure Based Identification of Nonlinear Systems 8589
Neetu Chikyal, Vasundhara, Chayan Bhar, Asutosh Kar, Mads Græsbøll Christensen

Uncertainty-Aware Robust Learning on Noisy Graphs 8594
Shuyi Chen, Kaize Ding, Shixiang Zhu

Diffusion Features to Bridge Domain Gap for Semantic Segmentation 8599
Yuxiang Ji, Boyong He, Chenyuan Qu, Zhuoyue Tan, Chuan Qin, Liaoni Wu

Noise-Robust Speech Emotion Recognition Using Shared Self-Supervised Representations with Integrated Speech Enhancement 8604
Jing-Tong Tzeng, Seong-Gyun Leem, Ali N. Salman, Chi-Chun Lee, Carlos Busso

Hazy Remote Sensing Image Semantic Segmentation with Weak Annotations Via Pre-Training Optimization and Co-Training 8609
Junda Xu, Libao Zhang

Quantum-Train with Tensor Network Mapping Model and Distributed Circuit Ansatz 8614
Chen-Yu Liu, Chu-Hsuan Abraham Lin, Kuan-Cheng Chen

CASC-XVC: Zero-Shot Cross-Lingual Voice Conversion with Content Accordant and Speaker Contrastive Losses 8618
Han-Jie Guo, Hui-Peng Du, Zheng-Yan Sheng, Li-Ping Chen, Yang Ai, Zhen-Hua Ling

MS-RainMamba: Learning Multi-Scale State Space Models for Single Image Deraining 8623
Haibo Li, Zhanshuo Liu, Tuo Zhao, Tingting Zhao, Yarui Chen, Ning Xie

SPED:A Sight-Singing Dataset for Performance Evaluation 8628
Yan Zhang, Jie Luo, Tianrui Li, Wei Xu

Cooperative Neural Radiance Field for Dynamic Scene Deblurring 8633
Dong Liu, Zhiyong Wang, Linlin Guo

Beyond Jensen's Inequality: Speeding Up ML Estimation of Generalized Hyperbolic Distributions 8638
Chenyu Gao, Ziping Zhao

HDA-GS: Hierarchical Density-Controlled for Anisotropic 3D Gaussian Splatting 8643
Zhanke Wang, Guanhua Wu, Zhiyan Wang, Lu Xiao, Runling Liu, Jiahao Wu, Ronggang Wang

Quickest Change Detection of Unknown Mean-Shifts Using the James-Stein Estimator 8648
Topi Halme, Venugopal V. Veeravalli, Visa Koivunen

CompMTL: Layer-Wise Competitive Multi-Task Learning 8653
Tiancong Cheng, Ying Zhang, Rajiv Ratn Shah, Roger Zimmermann, Zhiwen Yu, Bin Guo

Performance Bounds for Single-Bit AOA-Based RF Source Localization 8658
Shaunak Kubal, Anastasia Lavrenko, André Kokkeler

InjectTST: Injecting Global Information into Independent Channels for Long Time Series Forecasting 8663
Ce Chi, Xing Wang, Kexin Yang, Zhiyan Song, Di Jin, Lin Zhu, Chao Deng, Junlan Feng

Audio-Visual Deepfake Detection with Local Temporal Inconsistencies 8668
Marcella Astrid, Enjie Ghorbel, Djamila Aouada

L3D-Pose: Lifting Pose for 3D Avatars from a Single Camera in the Wild 8673
Soumyaratna Debnath, Harish Katti, Shashikant Verma, Shanmuganathan Raman

Hierarchical Nash Equilibrium Over Variational Equilibria Via Fixed-Point Set Expression of Quasi-Nonexpansive Operator 8678
Shota Matsuo, Keita Kume, Isao Yamada

Class-Wise Adaptive Logits Distillation with Meta-Learning 8683
Xiao Huang, Wu Chen, Wei Zhou

FreeAlign: Superior Text-Image Alignment by Modulating Prompt Attention 8688
Yibo Zhang, Dahua Gao, Feng Xie, Minxi Yang, Wenlong Wang, Ruichao Liu

FusionClassNet: A Multi-Scale Feature Fusion Network with Contrastive Loss-Driven Classification for Enhanced Lung Tumor Image Representations..... 8693
Hong Liu, Chengchuang Lin, Lewen Nie, Yonglin Peng, Qizhi Zhang, Gansen Zhao

Investigating Training Objectives for Generative Speech Enhancement..... 8698
Julius Richter, Danilo De Oliveira, Timo Gerkmann

Multi-Task Joint 3D Swin Transformer Learning for Segmentation and Classification of Hyperspectral Medicine Images 8703
Dong Zhang, Meijun Sun

FEA-DETR: An Enhanced ConvNet for Detecting Prohibited Objects in X-Ray Images Using Frequency and Edge Aware Information 8708
Shilong Hong, Yanzhou Zhou, Weichao Xu

Language-Based Audio Moment Retrieval 8713
Hokuto Munakata, Taichi Nishimura, Shota Nakada, Tatsuya Komatsu

Label Relationship Graph-Enhanced Class Hierarchy for Incremental Classification of Remote Sensing Images..... 8718
Yang Chu, Yuntao Qian

A Federated Learning-Based Intrusion Detection System for Satellite-Terrestrial Integrated Networks 8723
Mengke Wan, Jiang Fang, Chen Guo, Liru Geng, Yinlong Liu, Wei Ma, Chao Xu, Mohan Su

Efficient Non-Sequential Relational Modeling for Temporal Knowledge Graph Link Predictions 8728
Linhua Dong, Xiaofei Zhou, Qiannan Zhu, Xu Bai, Gang Xiong

Stable Control Visual AutoRegressive Model: Precise and Efficient Image Generation Via Scale Alignment..... 8733
Feng Xie, Dahua Gao, Ruichao Liu, Minxi Yang, Yibo Zhang, Wenlong Wang

PNetGPT: Proprietary Protocol Network Traffic Generation with Pre-Trained Transformer 8738
Zedong Li, Shijie Li, Dongliang Fang, Xin Chen, Zhanwei Song, Zhi Li, Shichao Lv, Limin Sun

Enhancing Image Editing with Chain-Of-Thought Reasoning and Multimodal Large Language Models..... 8743
Mengxue Kang, Xinyu Zhang, Fei Wei, Shuang Xu, Yuhe Liu

Deep Learning Amplified Early Stopping Bias: Overestimating Performance on Small Datasets..... 8748
Nona Rajabi, Antônio H. Ribeiro, Miguel Vasco, Danica Kragic

Exploring Local Interpretable Model-Agnostic Explanations for Speech Emotion Recognition with Distribution-Shift..... 8753
Maja J. Hjuler, Line H. Clemmensen, Sneha Das

SepNet: Deep Convolutional Neural Network for Specific Emitter Identification with High Accuracy..... 8758
Rong Wang, Xu Zhuang, Weixi Zhou, Zhicheng Dong

Attention Weighting and Conditional Entropy-Driven Quantization Loss for Neural Audio Codecs 8763
Huiyu Zhang, Weiping Tu, Yuhong Yang, Xinmeng Xu, Yanzhen Ren, Rong Zhu

Improving Embeddings by Refining Meanings for Temporal Knowledge Graph Link Predictions 8768
Linhua Dong, Xiaofei Zhou, Qiannan Zhu, Xu Bai, Gang Xiong

AccentBox: Towards High-Fidelity Zero-Shot Accent Generation 8773
Jinzuomu Zhong, Korin Richmond, Zhiba Su, Siqi Sun

A Two-Stage AIGC Image Quality Assessment with T2I Correspondence and Visual Perception 8778
Jili Xia, Lihuo He, Bo Hu, Bo Han, Xinbo Gao

Mix-Mask Augmentation and Self-Reconstruction for Cross-Domain Few-Shot Hyperspectral Image Classification 8783
Qin Xu, Jie Wei, Qihang Wu, Jiahui Wang, Xiao Wang, Jinpei Liu, Bo Jiang

Quantum Multimodal Contrastive Learning Framework 8788
Chi-Sheng Chen, Aidan Hung-Wen Tsai, Sheng-Chieh Huang

Multi-Agent Hierarchical Graph Attention Actor-Critic Reinforcement Learning 8793
Tongyue Li, Dianxi Shi, Songchang Jin, Zhen Wang, Huanhuan Yang, Yang Chen

RobNAS: Robust Neural Architecture Search for Point Cloud Adversarial Defense 8798
Shuoyang Sun, Kaiwen Zhang, Hao Fang, Bin Chen, Jiawei Li, Enze Huo, Shu-Tao Xia

Music Tagging with Classifier Group Chains 8803
Takuya Hasumi, Tatsuya Komatsu, Yusuke Fujita

TDOA Localization Via Fixed-Point Iteration 8808
Yanbin Zou, Yangpeng Xiao, Zekai Zhang, Huaping Liu

Passive Non-Line-Of-Sight Imaging with Parallel Encoder 8813
Xiaolong Du, Ruixu Geng, Jiarui Zhang, Yan Chen, Yang Hu

Selective Consistency Gradient Attack: Resolving Multi-Target Gradient Conflicts in Object Detection 8818
Dong Huang, Tianrun Jia, Pengyu Zhang, Ruihang Ji, Shuzhi Sam Ge

Enhancing Prosody Transfer in Speech Synthesis by Using Prosodically-Aligned References 8823
Lin Liu

Hybrid Spatial-Frequency Attention Network for Fine-Grained Skeleton-Based Action Recognition 8828
Dandan Zhang, Sicong Zhan, Jia Wang, Zhiqing Guo

Collision-Less and Balanced Sampling for Language-Queried Audio Source Separation 8833
Binh Thien Nguyen, Daiki Takeuchi, Masahiro Yasuda, Daisuke Niizumi, Noboru Harada

MM-LogVec: System Log Anomaly Detection Method Based on Multimodal Representation Learning 8838
Jingwen Li, Ru Zhang, Jianyi Liu

Adapter-Based Multi-Agent AVSR Extension for Pre-Trained ASR Models 8843
Christopher Simic, Korbinian Riedhammer, Tobias Bocklet

BeatKAN: An Efficient and Drum-Attuned Beat Tracking Method Using Kolmogorov-Arnold Networks 8848
Zhicheng Zhang, Ganghui Ru, Wei Li

A Generative-Augmented Deep Matrix Factorization Model for POI Recommendations 8853
Chunyu Liu, Hongli Zhang, Zeshu Tian, JiaYu Zhang

Covert and Potent: A Weather-Camouflaged Backdoor Attacks on Self-Supervised Learning 8858
Yang Wei, Yonghao Yang, Bo Liu, Bin Xiao

IPNet: Interpretable Prototype Network for Multi-Source Domain Adaptation 8863
Rui Chen, Haifeng Xia, Siyu Xia, Ming Shao, Zhengming Ding

Basis Function Learning for Variable-Length and Continuous-Indexed Signals 8868
Siyuan Li, Lei Cheng, Feng Yin, Jianlong Li, Peter Gerstoft

N3C: Towards Replay-Based Novelty Continual Clustering with Class-Overlapping 8873
Yan Zhang, Guoqiang Wu, Bingzheng Wang, Teng Pang, Yilong Yin

MHGNet: Multi-Heterogeneous Graph Neural Network for Traffic Prediction 8878
Mei Wu, Yiqian Lin, Tianfan Jiang, Wenchao Weng

Individual Fairness for Fuzzy C-Means Clustering 8883
Zhijing Yang, Boyang Yan, Junjie Zheng, Yiding Tang, Chuan Qian, Hui Zhang

Explainable Orthogonal Attention Networks for EEG-Based Analysis: Leveraging Disentangled Representations to Enhance Diagnosis 8888
Ailar Mahdizadeh, Puria Azadi Moghadam, Shahriar Mirabbasi, Panos Nasiopoulos

Improving Multimodal Human Pose Estimation by Adversarial Modality Enhancement† 8893
Jiangnan Xia, Qilong Wu, Yanyin Guo, Yi Li, Jianghan Cheng, Junwei Li, Zhiyuan Zhang

DACAT: Dual-Stream Adaptive Clip-Aware Time Modeling for Robust Online Surgical Phase Recognition 8898
Kaixiang Yang, Qiang Li, Zhiwei Wang

Investigating Factors Related to the Naturalness of Synthesized Unison Singing 8903
Kaito Nishizawa, Ryuichi Yamamoto, Wen-Chin Huang, Tomoki Toda

An Adversarial Perturbation Generation Method for Image Anti-Forensics Based on Dual-Path Spatial Attention GAN 8908
Yihong Lu, Jianyi Liu, Ru Zhang

Improving Contextual ASR with Enhanced Phrase-Level Representation Based on MCTC Loss 8913
Ming Fang, Tao Wei, Kai Guo, Ziyang Zhuang, Yan Shi, Ning Cheng, Shaojun Wang, Jing Xiao

Mean-Field Aided QMIX: A Scalable and Flexible Q-Learning Approach for Large-Scale Agent Groups 8918
Enze Zhang, Huaze Tang, Xiao-Ping Zhang, Wenbo Ding

Adaptive-Similarity-Based Brain Dynamic Functional Connectivity with Spatial-Temporal Attention and Domain Adaptation for Schizophrenia Diagnosis 8923
Yixin Ji, Vince D. Calhoun, Rongtao Jiang, Daoqiang Zhang, Shile Qi

RevPv8: Reverse Information for Road Space and Lane Line Segmentation in Highway Surveillance Scene 8928
Jingyi Wu, Jun Jing, Zeyong Zhao, Yunfeng Kang, Shiyu Zhang, Delan Kong, Wei Tang, Peng Jiang

Adaptive Layered-Trust Robust Defense Mechanism for Personalized Federated Learning 8933
He Wang, Zhen Xu, Yan Zhang, Yu Wang

Self-Prompting Driven SAM2 for 3D Medical Image Segmentation 8938
Sheng Wei, Song Qiu, Mei Zhou, He Zhang, Yan Wang, Qingli Li

Bridge-SR: Schrödinger Bridge for Efficient SR 8943
Chang Li, Zehua Chen, Fan Bao, Jun Zhu

Improving Open-Ended Referring Expression Comprehension Via Dual-Language Constraints 8948
Chenglong Lei, Chao Yang, Jianbo Zheng

AnCoGen: Analysis, Control and Generation of Speech with a Masked Autoencoder 8953
Samir Sadok, Simon Leglaive, Laurent Girin, Gaël Richard, Xavier Alameda-Pineda

ASR Benchmarking: Need for a More Representative Conversational Dataset 8958
Gaurav Maheshwari, Dmitry Ivanov, Théo Johannet, Kevin El Haddad

Enhancing Fairness in Gaussian Mixture Clustering Through Impact Factor 8963
Zhijing Yang, Chuan Qian, Junjie Zheng, Yiding Tang, Boyang Yan, Hui Zhang

Device Selection for Resource-Efficient Edge Caching in a Federated Learning Framework 8968
Deepali Kushwaha, Meenal Narkhede, Archana Limaye, Niranjan Pol, Rajesh M. Hegde

Zero-Shot Micro-Video Classification with Dual Alignment Topic Model 8973
Jialong Wang, Shilong Zhang, Zhiguo Gong

Knowledge Distillation Based Training of Unified Conformer CTC Models for Multi-Form ASR 8978
Takashi Fukuda, Gakuto Kurata, George Saon

SEHAP: Secure and Efficient Handover Authentication Protocol in LEO Satellite Non-Terrestrial Networks 8983
Yunchuan Guo, Jing Wang, Kui Geng, Zifu Li, Fenghua Li, Liang Fang

SFADNet: Spatio-Temporal Fused Graph Based on Attention Decoupling Network for Traffic Prediction 8988
Mei Wu, Wenchao Weng, Jun Li, Yiqian Lin, Jing Chen, Dewen Seng

OTMEA : Multi-Modal Entity Alignment Via Optimal Transport 8993
Cunda Wang, Weihua Wang, Xinyu Li, Qiuyu Liang, Feilong Bao

M2F2Net: Multi-Stage Mixed Feature Fusion Network for Remote Sensing Change Detection 8998
Binhao Gu, Lei Song, Youyong Kong, Binjie Gu

Integrating Concept Associations for Query Focused Knowledge Summarization 9003
Jian Wang, Zhi Liu, Yuqing Sun, Xin Li

Toward Visual Pronunciation Learning: A Speech-To-Articulatory Animation Pipeline Leveraging Wav2vec 2.0 and rtMRI Landmarks 9008
Mushaffa Rasyid Ridha, Shinobu Hasegawa, Sakriani Sakti

Dual-Path Model for Pulmonary Artery Segmentation 9013
Lu Shen, Yingwen Chen, Changjian Wang, Zhengbo Zhang, Shuyi Zhou

Identifying Bots on Social Media Through Coordinated Group Perception 9018
Boyu Qiao, Kun Li, Wei Zhou, Shilong Li, Qianqian Lu, Songlin Hu

Standard-Essential Patent Prediction: Discussion of Patent Standardization Time 9023
Weidong Liu, Zonghan Bai, Keqin Gan, Yan Cao

Prompt-To-Correct: Automated Test-Time Pronunciation Correction with Voice Prompts 9028
Ayan Kashyap, Neil Shah, Vineet Gandhi

Collusion-Resistant Black-Box Watermarking in Federated Learning Through Weight Relevance Analysis 9033
Elena Rodríguez-Lois, Fernando Pérez-González

Token-Level Contextual Network with Ladder-Shaped Attention for End-To-End ASR 9038
Ming Fang, Kai Guo, Tao Wei, Ziyang Zhuang, Yan Shi, Ning Cheng, Shaojun Wang, Jing Xiao

SUGAR: Leveraging Contextual Confidence for Smarter Retrieval 9043
Hanna Zubkova, Ji-Hoon Park, Seong-Whan Lee

MonoIR: Inpainting and Reconstruction for Monocular Endoscope Deformation Scenes 9048
Ziteng Zhang, Wenyu Li, Sidun Liu, Peng Qiao, Yong Dou

Better Exploiting Spatial Separability in Multichannel Speech Enhancement with an Align-And-Filter Network 9053
Ching-Hua Lee, Chouchang Yang, Yashas Malur Saidutta, Rakshith Sharma Srinivasa, Yilin Shen, Hongxia Jin

LocRef-Diffusion:Tuning-Free Layout and Appearance-Guided Generation 9058
Fan Deng, Yaguang Wu, Xinyang Yu, Xiangjun Huang, Jian Yang, Guangyu Yan, Qiang Xu

Complex Open Information Extraction with Heterogeneous Syntax Forests 9063
Meishan Zhang, Peiming Guo, Hao Fei, Min Zhang

Input Uncertainty Attribution by Uncertainty Propagation 9068
Benedikt Kantz, Sophie Steger, Clemens Staudinger, Christoph Feilmayr, Johann Wachlmayr, Alexander Haberl, Stefan Schuster, Franz Pernkopf

Advancing NAM-To-Speech Conversion with Novel Methods and the MultiNAM Dataset 9073
Neil Shah, Shirish Karande, Vineet Gandhi

ICAA-Mamba: Vision Mamba for Image Color Aesthetics Assessment 9078
Qinghao Xie, Wenchao Jiang, Zhihua Wang, Qiuping Jiang

Non-Autoregressive Image Captioning with Multi-Label Classification and Self-Critical Sequence Training 9083
Yuanqiu Liu, Hong Yu, Hui Li, Xiaotong Zhang, Xin Han, Han Liu

CardioFlow: Learning to Generate ECG from PPG with Rectified Flow 9088
Yuta Nambu, Masahiro Kohjima, Ryuji Yamamoto

Deep Time Series Anomaly Detection with Local Temporal Pattern Learning 9093
Yizhou Li, Yijie Wang, Hongzuo Xu, Xiaohui Zhou

Enhancing DETR Efficiency with Inter-Object Relationship and Semantic Spectral Decomposition-Based Distillation 9098
Ziyu Huang

ValSub: Subsampling Validation Data to Mitigate Forgetting During ASR Personalization 9103
Haaris Mehmood, Karthikeyan Saravanan, Pablo Peso Parada, David Tuckey, Mete Ozay, Gil Ho Lee, Jungin Lee, Seokyeong Jung

GraphVCM: Virtual Center Mixing with Distance-Aware Regulation for Class Imbalanced Node Classification 9108
Yixiao Ren, Yunfei Han, Yi Wang, Zhengdong Luo, Jinlong Liu, Yupeng Ma

Single-Channel Distance-Based Source Separation for Mobile GPU in Outdoor and Indoor Environments 9113
Hanbin Bae, Byungjun Kang, Jiwon Kim, Jae-Yong Hwang, Hosang Sung, Hoon-Young Cho

Through-The-Wall Multi-Person Localization Using Translation and Rotation Synthetic Aperture Radar 9118
S. Sinha, A. Deshwal, A. Azizi, D. Pandey, N. Mehrotra, A. Pal, A. Sabharwal

Spectral-Aware Low-Rank Adaptation for Speaker Verification 9123
Zhe Li, Man-Wai Mak, Mert Pilanci, Hung-Yi Lee, Helen Meng

"I've Heard of You!": Generate Spoken Named Entity Recognition Data for Unseen Entities 9128
Jiawei Yu, Xiang Geng, Yuang Li, Mengxin Ren, Wei Tang, Jiahuan Li, Zhibin Lan, Min Zhang, Hao Yang, Shujian Huang, Jinsong Su

Chained Motion Vector Prediction for Video Coding 9133
Yoshitaka Kidani, Haruhisa Kato, Kei Kawamura

Learning from Reconstruction: A Two-Stage Global-To-Local Framework for Temporal Knowledge Graph Completion 9138
Wenjie Xu, Kai Liu, Zihao Jiang, Mengting Song, Boyi Zhang, Min Peng

Frequency-Enhanced Comprehensive Dependency Attention for Time Series Anomaly Detection 9143
Haonan Chen, Hongzuo Xu, Songlei Jian, Ruyi Zhang, Xingming Li, Zibo Yi

MusicLIME: Explainable Multimodal Music Understanding 9148
Theodoros Sotirou, Vassilis Lyberatos, Orfeas Menis Mastromichalakis, Giorgos Stamou

Microphone Array Beamforming for Speech Enhancement Based on Dynamic Mode Decomposition 9153
Wei Liu, Gongping Huang, Xin Liu, Jilu Jin, Jingdong Chen, Jacob Benesty

Next-Generation ANC: Integrating Dynamic Fixed-Filter Strategies with Extended Kalman Filtering for Enhanced Noise Suppression 9158
Fareedha, Vasundhara, Asutosh Kar, Mads Græsbøll Christensen

Conditional-Balanced Adversarial Delta Tuning for Cross-Domain Implicit Discourse Relation Recognition 9163
Chang Liu, HaoWen Sun, Haodong Zhao, Ruifang He, Xu He, Bo Wang

Semi-Automatic Labeling for Action Recognition by Diversity Preserving Sampling 9168
Ryuhei Ando, Takashi Shibata, Toru Takahashi

Aligned Contrastive Learning for Text-To-Music Retrieval 9173
Tatsuya Komatsu, Hokuto Munakata, Takuya Hasumi, Yusuke Fujita

Positive Enhanced Preference Alignment for Text-To-Image Models 9178
Haoyuan Sun, Bo Xia, Yifei Zhao, Yongzhe Chang, Xueqian Wang

Subjective Voice Quality of the IVAS Codec 9183
Anssi Rämö, Henri Toukomaa

DFingerNet: Noise-Adaptive Speech Enhancement for Hearing Aids 9188
Iosif Tsangko, Andreas Triantafyllopoulos, Michael Müller, Hendrik Schröter, Björn W. Schuller

DrawSpeech: Expressive Speech Synthesis Using Prosodic Sketches as Control Conditions 9193
Weidong Chen, Shan Yang, Guangzhi Li, Xixin Wu

MLSDET: Multi-LLM Statistical Deep Ensemble for Chinese AI-Generated Text Detection 9198
Dianhui Mao, Denghui Zhang, Ao Zhang, Zhihua Zhao

GradPFL: Gradient-Driven Adaptive Clustering in Personalized Federated Learning 9203
Shiyu Song, Hao Zheng, Zhigang Hu, Meiguang Zheng, Liu Yang, Aikun Xu

Enhancing Unsupervised Acoustic Word Embedding with Visual-Grounded Speech Model and Novel Word-Level ABX Evaluation Schemes 9208
Mau Nguyen, Shinobu Hasegawa, Sakriani Sakti

Task Vector Arithmetic for Low-Resource ASR 9213
Haruki Nagasawa, Shinta Otake, Shinji Iwata

VALL-T: Decoder-Only Generative Transducer for Robust and Decoding-Controllable Text-To-Speech 9218
Chenpeng Du, Yiwei Guo, Hankun Wang, Yifan Yang, Zhikang Niu, Shuai Wang, Hui Zhang, Xie Chen, Kai Yu

DPC: Large Model Alignment Method Based on Decoding Probability Correction 9223
Yanyi Huang, Yuying Liu, Lei Tian, Yue Zhang, Xuechen Zhao, Bin Zhou, Guodong Ma

Frequency-Domain Popularity Forecasting with Shape-Based Retrieval 9228
Canhua Guan, Zongxia Xie, Haoyu Wang, Haoyu Xing

SCBot: Building Lightweight and Flexible C&C Based on Smart Contract 9233
Chaoge Liu, Zhi Wang, Jie Yin, Yinsheng Liu, Kai Mao, Zhe Huang, Chumeng Deng

Hierarchical Perceptual Distillation Network for Lightweight Image Super-Resolution Reconstruction 9238
Qingting Tang, Zhiqing Guo, Liejun Wang

Self-Distillation Prototypes Network: Learning Robust Speaker Representations Without Supervision 9243
Yafeng Chen, Siqi Zheng, Hui Wang, Luyao Cheng, Qian Chen, Chong Deng, Shiliang Zhang, Wen Wang

Generative Model Based Optical Response Prediction for Plasmonic Sensing 9248
Anish Datta, Soma Bandyopadhyay, Subhasri Chatterjee, Tapas Chakravarty, Arpan Pal

Decoding Brain Structure and Gene Expression Interactions in Alzheimer's Disease Pathology 9253
Amashi Niwarthana, Chockalingam Kasi, Yi Hao Chan, Conghao Wang, Jagath C Rajapakse

Big-Moe: Bypassing Isolated Gating for Generalized Multimodal Face Anti-Spoofing 9258
Yingjie Ma, Zitong Yu, Xun Lin, Weicheng Xie, Linlin Shen

Boolean Matrix Tri-Factorization 9263
Christos Kolomvakis, Arnaud Vandaele, Nicolas Gillis

Few-Shot Keyword-Incremental Learning Using Compositional Information 9268
Ilseok Kim, Ju-Seok Seong, Joon-Hyuk Chang

Graph Contrastive Learning with Decoupled Augmentation 9273
Shihao Gao, Caoshuo Li, Cunli Mao, Xulong Zhang, Xiaoyang Qu, Taisong Jin, Jianzong Wang

Step-By-Step Correction of LLM-Based Math Word Problems Solutions 9278
Yiyao Li, Dhanish Musharraf Ubaidali, Lu Wang, Wenyu Zhang

AP-Net: Semi-Supervised Ultrasound Cardiac Segmentation Using Enhanced Anatomical Prior 9283
Yuhuan Lu, Jintang Li, Jianxin Lin, Ying Yuan, Jagath C. Rajapakse, Ningbo Zhu, Chunlian Wang, Kenli Li

Low Complexity Super Resolution for Resampling-Based Video Coding 9288
Chaoyi Lin, Yue Li, Junru Li, Kai Zhang, Li Zhang

Out-Of-Distribution Radar Detection in Compound Clutter and Thermal Noise Through Variational Autoencoders 9293
Y. A. Rouzoumka, E. Terreaux, C. Morisseau, J.-P. Ovarlez, C. Ren

Variance-Covariance Regularization for Improved End-To-End Diarization 9298
Lahiru Samarakoon, Samuel J. Broughton, Ivan Fung

PRESS: Defending Privacy in Retrieval-Augmented Generation Via Embedding Space Shifting 9303
Jiaming He, Cheng Liu, Guanyu Hou, Wenbo Jiang, Jiachen Li

KAN-HyperpointNet for Point Cloud Sequence-Based 3D Human Action Recognition 9308
Zhaoyu Chen, Xing Li, Qian Huang, Qiang Geng, Tianjin Yang, Shihao Han

Understanding Neural Networks in Profiled Side-Channel Analysis 9313
Yimeng Chen, Bo Wang, Changshan Su, Ao Li, Gen Li, Yuxing Tang

Joint-Wise Distributed Perception Graph Convolutional Network for Skeleton-Based Action Recognition 9318
Qian Huang, Qiang Geng, Zhaoyu Chen, Xin Li, Yangyang Li, Xing Li

Multimodal Emotion Recognition with Target Speaker-Based Facial Embeddings 9323
Serin Heo, Jehyun Kyung, Joon-Hyuk Chang

Time-Varying EEG Signal Reconstruction Based on Local Graph Signal Smoothness 9328
Yuting Cao, Jian Wang, Zhengpin Li, Xue-Qin Jiang, Xinxin Hou

Frank-Wolfe Method with Proximal Regularization for Constrained Federated Learning with Non-Iid Data 9333
Robin Francis, Sundeep Prabhakar Chepuri

AER-LLM: Ambiguity-Aware Emotion Recognition Leveraging Large Language Models 9338
Xin Hong, Yuan Gong, Vidhyasaharan Sethu, Ting Dang

CMFNThinker: A Novel Cross-Source Multi-Modal Fake News Detection Model 9343
Kaijia Tian, Guozheng Rao, Xin Wang, Mufan Yu, Jiayin Zhang, Li Zhang

Differentially Private and Communication-Efficient Decentralized Learning Using Deep Quantizers 9348
Robin Francis, Sundeep Prabhakar Chepuri

PPDformer: Channel-Specific Periodic Patch Division for Time Series Forecasting 9353
Meng Wan, Qi Su, Huan Hao, Jue Wang, Yuexiu Cui, Yuxuan Bi, Rongqiang Cao, Peng Shi, Yangang Wang, Zonghua Qiu, Zongshan Zhang

Text-Guided Editable 3D City Scene Generation 9358
Yuchuan Feng, Jihang Jiang, Jie Ren, Wenrui Li, Ruotong Li, Xiaopeng Fan

Simple Contrastive Learning with Knowledge Graphs for Story Generation 9363
Yu Zhu, Rong Pan

Evaluation of Deep Audio Representations for Hearables 9368
Fabian Gröger, Pascal Baumann, Ludovic Amruthalingam, Laurent Simon, Ruksana Giurda, Simone Lionetti

Mask Augmented Object-Centric Contrastive Learning for Amodal Instance Segmentation 9373
Tomokazu Kaneko, Ryosuke Sakai, Takashi Shibata, Soma Shiraishi

Learned ReLU-Based Soft Thresholding: A Data-Driven Method for Non-Negative Sparse Signal Recovery 9378
Akash Sen, Pradyumna Pradhan, Ramunaidu Randhi, C. S. Sastry

FDR Control for Complex-Valued Data with Application in Single Snapshot Multi-Source Detection and DOA Estimation 9383
Fabian Scheidt, Jasin Machkour, Michael Muma

GateM2Former: Gated Feature Selection and Expert Modeling in Multimodal Emotion Recognition 9388
Weixiang Xu, Zhongren Dong, Runming Wang, Xinzhou Xu, Zixing Zhang

DetailTTS: Learning Residual Detail Information for Zero-Shot Text-To-Speech 9393
Cong Wang, Yichen Han, Yizhong Geng, Yingming Gao, Fengping Wang, Bingsong Bai, Qifei Li, Jinlong Xue, Yayue Deng, Zhengqi Wen, Ya Li

SlimSpeech: Lightweight and Efficient Text-To-Speech with Slim Rectified Flow 9398
Kaidi Wang, Wenhao Guan, Shenghui Lu, Jianglong Yao, Lin Li, Qingyang Hong

An Efficient Hybrid Quantum Variational Classifier with Matrix Product State 9403
Wanqi Sun, Jungang Xu, Chenghua Duan

PersoDA: Personalized Data Augmentation for Personalized ASR 9408
Pablo Peso Parada, Spyros Fontalis, Md Asif Jalal, Karthikeyan Saravanan, Anastasios Drosou, Mete Ozay, Gil Ho Lee, Jungin Lee, Seokyeong Jung

Efficient Quality Controllable Neural Image Compression Based on QD-Model 9413
Shaokang Wang, Guoqing Xiang, Jinchang Xu, Shanghang Zhang, Xiaodong Xie

Radio Map Estimation Via Latent-Domain Plug-And-Play Denoisers 9418
Le Xu, Lei Cheng, Junting Chen, Wenqiang Pu, Xiao Fu

Enhanced Sparse Bayesian Learning Methods with Application to Massive MIMO Channel Estimation 9423
Arttu Arjas, Italo Atzeni

Intent-Driven In-Context Learning for Few-Shot Dialogue State Tracking 9428
Zihao Yi, Zhe Xu, Ying Shen

Multi-Modal Dynamic Point Cloud Geometric Compression Based on Bidirectional Recurrent Scene Flow* 9433
Fangzhe Nan, Frederick Li, Zhuoyue Wang, Gary K.L. Tam, Zhaoyi Jiang, DongZheng DongZheng, Bailin Yang

Dynamic-Static Feature Fusion with Multi-Scale Attention for Continuous Blood Glucose Prediction 9438
Jing Gao, Chenhua Guo, Yingshu Liu, Peng Li, Jianing Zhang, Meng Liu

MSA-ASR: Efficient Multilingual Speaker Attribution with Frozen ASR Models 9443
Thai-Binh Nguyen, Alexander Waibel

Uncertainty Prediction for Prominence Classification with Chroma Features 9448
Julian Linke, Sophie Steger, Philipp Steinwender, Gernot Kubin, Franz Pernkopf, Barbara Schuppler

Correlative3D: Inter-Object Correlation-Aware 3D Scene Understanding 9453
Tingxuan Gao, Wenming Yang, Yang Wu, Yehu Shen

Deformable Attention-Based Edge-Aware Network for Single Image Super-Resolution 9458
Ju Zhang, Baojiang Zhong, Kai-Kuang Ma

A Self-Evolving Framework for Multi-Agent Medical Consultation Based on Large Language Models 9463
Kai Chen, Ji Qi, Jing Huo, Pinzhuo Tian, Fanyu Meng, Xi Yang, Yang Gao

PD-VOST: Parkinson's Disease Voice Spectrogram Transformer 9468
Ilias Tougui, Mehdi Zakroum, Ouassim Karrakchou, Mounir Ghogho

KNN-CL: Enhancing Continual Learning with Nearest Neighbor Retrieval 9473
Enzhi Wang, Qicheng Li, Hao Chen, Ruiqi Sun, Xin Zhou

Contrastive Lyrics Alignment with a Timestamp-Informed Loss 9478
Timon Kick, Florian Grötschla, Luca A. Lanzendörfer, Roger Wattenhofer

Learning Weighted Least Squares Data Term for Poisson Image Deconvolution 9483
Abhijit Singh, Emmanuel Soubies, Caroline Chaux

Generalizable Audio Deepfake Detection Via Latent Space Refinement and Augmentation 9488
Wen Huang, Yanmei Gu, Zhiming Wang, Huijia Zhu, Yanmin Qian

RAPID: Recognition of Any-Possible DrIver Distraction Via Multi-View Pose Generation Models 9493
Jingyu Lei, Shengyu Hao, Gaoang Wang, Der-Horng Lee

Multimodal ARMAX Model for Characterization of Human Body System 9498
Huiling Li, Qian He, Zhao Jin

Multi-View Subspace Classification: A Hierarchical Contrastive Approach and Low-Rank Latent Representation 9503
Deyu Zeng, Tengyu Zhang, Zongze Wu, Wei Liu, Chris Ding

Addressing Speed-Induced Dispersion in Stepped-Frequency PMCW Radar Systems 9508
Moritz Kahlert, Tai Fei, Claas Tebruegge, Shunqiao Sun, Markus Gardill

D3RM: A Discrete Denoising Diffusion Refinement Model for Piano Transcription 9512
Hounsu Kim, Taegyun Kwon, Juhan Nam

Bootstrapping LLM-Based Fact-Checking Via Iterative Rationalization Finetuning 9517
Xiucheng Lyu, Chengyu Cao, Mingwei Sun, Bin Liang, Liang Yao, Ruifeng Xu

PhysID: Physics-Based Interactive Dynamics from a Single-View Image 9522
Sourabh Vasant Gothe, Ayon Chattopadhyay, Gunturi Venkata Sai Phani Kiran, Pratik, Vibhav Agarwal, Jayesh Rajkumar Vachhani, Sourav Ghosh, Parameswaranath VM, Barath Raj KR

Dense Point Clouds Matter: Dust-GS for Scene Reconstruction from Sparse Viewpoints 9527
Shen Chen, Jiale Zhou, Lei Li

Hyperedge Representations with Hypergraph Wavelets: Applications to Spatial Transcriptomics 9532
Xingzhi Sun, Charles Xu, João F. Rocha, Chen Liu, Benjamin Hollander-Bodie, Laney Goldman, Marcello DiStasio, Michael Perlmutter, Smita Krishnaswamy

V-Phanton:Voltage-Based Physically-Triggered Backdoor Attack Against Facial Recognition 9537
Yan Jiang, Ruishan Li, Yushi Cheng, Xiaoyu Ji, Wenyuan Xu

Accelerating Codec-Based Speech Synthesis with Multi-Token Prediction and Speculative Decoding 9542
Tan Dat Nguyen, Ji-Hoon Kim, Jeongsoo Choi, Shukjae Choi, Jinseok Park, Younglo Lee, Joon Son Chung

DOSE: Drum One-Shot Extraction from Music Mixture 9547
Suntae Hwang, Seonghyeon Kang, Kyungsu Kim, Semin Ahn, Kyogu Lee

A Dynamical Equation Approach for Quasi-Periodic Gaussian Processes 9552
Unnati Nigam, Radhendushka Srivastava, Michael Burke, Faezeh Marzbanrad

YOLO-KED: A Novel Framework for Rotated Object Detection in Complex Environments 9557
Zhaoyu Zhuang, Penglei Liu, Dejia Xu, Jun Cheng

Anchor-Prompt-Based Segmentation and Embedding Model 9562
Shuman Li, Zhipeng Lin, Haotian Wang, Wenjing Yang, Hengzhu Liu

Meta-Learning for Finger Vein Recognition in Internet of Things Smart Home Security 9567
Hengyi Ren, Lijuan Sun, Jinting Ren, Xing Li, Ying Cao

VoiceDiT: Dual-Condition Diffusion Transformer for Environment-Aware Speech Synthesis 9572
Jaemin Jung, Junseok Ahn, Chaeyoung Jung, Tan Dat Nguyen, Youngjoon Jang, Joon Son Chung

CA-MHFA: A Context-Aware Multi-Head Factorized Attentive Pooling for SSL-Based Speaker Verification 9577
Junyi Peng, Ladislav Mošner, Lin Zhang, Oldřich Plchot, Themos Stafylakis, Lukáš Burget, Jan Černocký

Automatic Adaption of the Step Size in Gradient Descent Training 9582
Albino Nogueiras-Rodríguez, Ignasi Nogueiras-Marco

LLM-GAN: Constructing Generative Adversarial Network Through Large Language Models for Explainable Fake News Detection 9587
Yifeng Wang, Zhouhong Gu, Siwei Zhang, Suhang Zheng, Tao Wang, Tianyu Li, Hongwei Feng, Yanghua Xiao

GEE Maximization in UAV-Aided Mobile IoT Networks Using Deep Reinforcement Learning 9592
Aditya Singh, Rajesh M. Hegde

Emo-DPO: Controllable Emotional Speech Synthesis Through Direct Preference Optimization 9597
Xiaoxue Gao, Chen Zhang, Yiming Chen, Huayun Zhang, Nancy F. Chen

Non-Contact Quickest Abnormal Heart Rate Detection Using MIMO Radar 9602
Peichao Wang, Qian He

Progressive Subband Modeling for Artifacts-Free Speech Super-Resolution 9607
Donghyun Kim, Joon-Hyuk Chang

Cross-Channel Unlabeled Sensing Over a Union of Signal Subspaces 9612
Taulant Koka, Manolis C. Tsakiris, Benjamín Béjar Haro, Michael Muma

Boli: A Dataset for Understanding Stuttering Experience and Analyzing Stuttered Speech 9617
Ashita Batra, Mannas Narang, Neeraj Kumar Sharma, Pradip K. Das

Efficient Long Document Ranking Via Adaptive Token Pruning with Query-Document Alignment 9621
Beiya Dai, Meilin Chen, Peng Yu, Xinbing Wang, Chenghu Zhou, Zhouhan Lin

Improving Height Prediction for Vision-Based Roadside 3D Object Detection 9626
Tengfei Zhang, Heng Zhang, Rengang Li, Yaqian Zhao, Qi Deng, Ruyang Li

SpecViT: A Custom Vision-Transformer Based Approach for Audio Deepfake Detection 9631
Sharmistha Modak, Arnab Kumar Das, Ruchira Naskar

DVM: Towards Controllable LLM Agents in Social Deduction Games 9636
Zheng Zhang, Yihuai Lan, Yangsen Chen, Lei Wang, Xiang Wang, Hao Wang

Robust Fusion of Bone and Air-Conducted Sensors for Speech Enhancement with Adaptive Temporal-Frequency Attention 9641
Zhenglong Liu, Zhe Chen, Fuliang Yin

Coarse-To-Fine Text-To-Music Latent Diffusion 9646
Luca A. Lanzendörfer, Tongyu Lu, Nathanaël Perraudin, Dorien Herremans, Roger Wattenhofer

SSE: A Speaking Style Extractor Based on Fine-Grained Contrastive Learning Between Speech and Descriptive Text 9651
Zixing Zhang, Yimeng Wu, Zhongren Dong, Wulong Xiang, Shengfan Shen, Björn W. Schuller

A Fast Saturation Based Dehazing Framework with Accelerated Convolution and Attention Block 9656
Shuocheng Wang, Jiaming Liu, Yilian Zhong, Ruoxi Zhu, Jiazheng Lian, Hao Zhang, Yibo Fan

Enhancing Large Language Models on Domain-Specific Tasks: A Novel Training Strategy Via Domain Adaptation and Preference Alignment 9661
Jingyang Deng, Zeren Zhang, Jo-Ku Cheng, Jinwen Ma

Raw Audio Deep Learning Filter Banks for Acoustic Scene Classification 9666
Daniele Salvati

EEG Correlation Analysis-Guided Graph Local Enhanced Feature Learning for Emotion Recognition 9671
Xinhui Li, Guowang Zhuang, Minchao Wu, Zhao Lv

Incremental Disentanglement for Environment-Aware Zero-Shot Text-To-Speech Synthesis 9676
Ye-Xin Lu, Hui-Peng Du, Zheng-Yan Sheng, Yang Ai, Zhen-Hua Ling

A Modular-Based Strategy for Mitigating Gradient Conflicts in Simultaneous Speech Translation 9681
Xiaoqian Liu, Yangfan Du, Jianjin Wang, Yuan Ge, Chen Xu, Tong Xiao, Guocheng Chen, Jingbo Zhu

LEF-TTS: Lightweight and Efficient End-To-End Text-To-Speech Synthesis with Multi-Stream Generator 9686
Yan Shi, Jin Shi, Minchuan Chen, Chenfeng Miao, Ming Fang, Ning Cheng, Shaojun Wang, Jing Xiao

AdaPPA: Adaptive Position Pre-Fill Jailbreak Attack Approach Targeting LLMs 9691
Lijia Lv, Weigang Zhang, Xuehai Tang, Jie Wen, Feng Liu, Jizhong Han, Songlin Hu

MAEM: A Multi-Aspect Extraction Model for Enhanced Embedding in RAG 9696
Ningyuan Yi, Chen Liu, Yue Wang, Jianjun Yu

Sparse-View X-Ray 3D Reconstruction Using Hybrid Representation Neural Attenuation Fields 9701
Yanping Fu, Hao Geng, Zhuangzhuang Zhao, Shaojie Zhang, Haifeng Zhao

M3ADD: A Novel Benchmark for Physiology Signal-Based Automatic Depression Detection with Multimodal Multitask Multievent Framework 9706
Changzeng Fu, Kaifeng Su, Yikai Su, Fengkui Qian, Yixuan Zhang, Chaoran Liu, Siyang Song, Le Yang, Xiaoyong Lv, Peng Shan, Yuliang Zhao

Fading-Invariant Adversarial Attacks on Neural Modulation Recognition 9711
Xinze Zhang, Dengao Zhu, Xiyao Dong, Kun He

Less is More: Efficient Scene Graph Generation with Reparameterization 9716
Jonghwan Hong, Seonghyeok Noh, Bonhwa Ku, Hanseok Ko

MedCAM-OsteoCls: Medical Context Aware Multimodal Classification of Knee Osteoarthritis 9721
Akshay Daydar, Alik Pramanick, Arijit Sur, Subramani Kanagaraj

Addressing Pilot Contamination in Channel Estimation with Variational Autoencoders 9726
Amar Kasibovic, Benedikt Fesl, Michael Baur, Wolfgang Utschick

Underwater Image Restoration Via Polymorphic Large Kernel CNNs 9731
Xiaojiao Guo, Yihang Dong, Xuhang Chen, Weiwen Chen, Zimeng Li, FuChen Zheng, Chi-Man Pun

Learning Permutations in Monarch Factorization 9736
Mimoun Mohamed, Valentin Emiya, Caroline Chaux

Multiple Sclerosis Detection with Reinforcement Learning and Differential Evolution 9741
Jing Yang, Jin Yang, Chenwei Wu, Gaozhe Jiang, Yaning Lv, Bingyan Liu, Feng Xu

LLGS: Illuminating Gaussian Splatting Via Absorptance Modulation 9746
Jianwen Gan, Wenxin Li, Bo Zheng, Chengliang Wang, Yingbo Wu

Hybrid Feature Global Attention Network for Noisy-Reverberant Speech Enhancement 9751
Zehua Zhang, Shiyun Xu, Yinghan Cao, Changjun He

Dynamic Speech Generation to Enhance Intelligibility in Noisy Environments 9756
Olympia Simantiraki, Maria Markaki, Yannis Pantazis

Open-Amp: Synthetic Data Framework for Audio Effect Foundation Models 9761
Alec Wright, Alistair Carson, Lauri Juvela

Federated Prototype Guided Adaption for Vision-Language Models 9766
Youchao Liu, Dingjiang Huang

JSUnet: A New Hybrid U-Shaped Network for Jamming Suppression 9771
Shuang Li, Ganggang Dong

Gaze-Assisted Human-Centric Domain Adaptation for Cardiac Ultrasound Image Segmentation 9776
Ruiyi Li, Yuting He, Rongjun Ge, Chong Wang, Daoqiang Zhang, Yang Chen, Shuo Li

Stereo Downmix in 3GPP IVAS for EVS Compatibility 9781
Takehiro Moriya, Stéphane Ragot, Arnaud Lefort, Alexandre Guérin, Noboru Harada, Ryosuke Sugiura, Yutaka Kamamoto

LAVCap: LLM-Based Audio-Visual Captioning Using Optimal Transport 9786
Kyeongha Rho, Hyeongkeun Lee, Valentio Iverson, Joon Son Chung

ECBANet: Exploiting Complementary Information for Efficient Burst Super-Resolution 9791
Liwen Zhang, Dingyong Gou, Cong Li, Yanlin Wu, Changjiang Xie, Ke Ren, Zhe Xu

COAST: Contrastive Learning with Augmented Spatio-Temporal Encoding for Next POI Recommendation 9796
Bada Xin, Xin Wan, Zhuojun Jiang, Faqiang Liu, Su Chen, Rong Yang, Qingyun Liu

Minimizing Disparities Between Real and Pseudo Queries for Unsupervised Visual Grounding 9801
Hui Jiang, Changkai Ji, Jilan Xu, Yanhao Zhu, Yuejie Zhang, Rui Feng, Tao Zhang, Shang Gao

CardioRiskNet: Attention-Based CVAE-Enabled GCN for Risk Prediction in STEMI 9806
Akshat Gupta, Anubha Gupta, Manu Shetty, Dixit Goyal, Girish M. P., Mohit Gupta

DICS: Find Domain-Invariant and Class-Specific Features for Out-Of-Distribution Generalization 9811
Qiaowei Miao, Yawei Luo, Yi Yang

Graph-Enhanced Dual-Stream Feature Fusion with Pre-Trained Model for Acoustic Traffic Monitoring 9816
Shitong Fan, Feiyang Xiao, Wenbo Wang, Shuhan Qi, Qiaoxi Zhu, Wenwu Wang, Jian Guan

Unveiling Deepfakes with Latent Diffusion Counterfactual Explanations 9821
Chen Yang, Bo Peng, Jing Dong, Xiaoyu Zhang

HBRW: A Hardness-Based Re-Weighting Approach for Long-Tailed Medical Image Classification 9826
Yongheng Xu, Hanjiang Lai

Completing Sets of Prototype Transfer Functions for Subspace-Based Direction of Arrival Estimation of Multiple Speakers 9831
Daniel Fejgin, Simon Doclo

Improving Continuous Sign Language Recognition Via Cross-Frame Interactions in Expanded Contextual Spaces 9836
Yiheng Yu, Sheng Liu, Yuan Feng, Min Xu, Zhelun Jin, Xuhua Yang

Enhancing Whisper's Accuracy and Speed for Indian Languages Through Prompt-Tuning and Tokenization 9841
Kumud Tripathi, Raj Gothi, Pankaj Wasnik

Combined Object-Based Audio and MASA Format for Enhanced Spatial Mobile Communication 9846
Mikko-Ville Laitinen, Adriana Vasilache, Anssi Rämö, Jouni Paulus, Václav Eksler

Spatio-Temporal Mixed Graph Neural Controlled Differential Equations with Adaptive Connection Sampling for Irregular Multivariate Time Series Anomaly Detection 9851
Xudong Jia, Wei Peng, Chiran Shen, Baokang Zhao, Peng Xun

ReTD: Reconstruction-Based Traceability Detection for Generated Images 9856
Weizhuo Chen, Fangfang Yuan, Cong Cao, Kun Peng, Dakui Wang, Yanbing Liu

ZipEnhancer: Dual-Path Down-Up Sampling-Based Zipformer for Monaural Speech Enhancement 9861
Haoxu Wang, Biao Tian

OpenACE: An Open Benchmark for Evaluating Audio Coding Performance 9866
Jozef Coldenhoff, Niclas Granqvist, Milos Cernak

DecoupledSynth: Enhancing Zero-Shot Text-To-Speech Via Factors Decoupling 9871
Jingyuan Xing, Shuaiqi Chen, Xiangmin Xu, Xiaofen Xing

KMG-LL: Knowledge-Enhanced Multimodal Graph for Dialogue Generation 9876
Yuezhou Dong, Tao He, Qian Dong, Ke Qin

Cognitive Load Monitoring Via Earable Acoustic Sensing 9881
Jiatao Quan, Khaldoon Al-Naimi, Xijia Wei, Yang Liu, Fahim Kawsar, Alessandro Montanari, Ting Dang

Graph Anomaly Detection Via Multi-Scale Reconstruction of Graph Encoder-Decoder Networks 9886
Jiaqi Yu, Hong Yang, Peng Zhang

3D-Speaker-Toolkit: An Open-Source Toolkit for Multimodal Speaker Verification and Diarization 9891
Yafeng Chen, Siqi Zheng, Hui Wang, Luyao Cheng, Tinglong Zhu, Rongjie Huang, Chong Deng, Qian Chen, Shiliang Zhang, Wen Wang, Xihao Li

Estimating Instrument Spectral Response Functions Using Sparse Representations and Quadratic Envelopes 9896
Jihanne El Haouari, Marcus Carlsson, Jean-Yves Tourneret, Herwig Wendt, Jean-Michel Gaucel, Christelle Pittet

RAS-GNN: Reconstructing APT Attack Scenario Using Graph Neural Network 9901
Zhicheng Huang, Ping Wang

MoME: Mixture of Multi-Domain Experts for Multivariate Long-Term Series Forecasting 9906
Xinyu Li, Zhiheng Yang, Hao Xu, Yunqi Cai, Hong Lu, Xin Wang, Jin Zhao, Fenglin Qi, Jiajie Shen

Generating Vocals from Lyrics and Musical Accompaniment 9911
Georg Streich, Luca A. Lanzendörfer, Florian Grötschla, Roger Wattenhofer

Audio-Faces Intra-Frame Alignment with Graph Attention Networks for Active Speaker Detection 9916
Yongkang Yin, Xusheng Yang, Liming Liang, Xu Li, Yuexian Zou

Adaptive Skeleton Prompt Tuning for Cross-Dataset 3D Human Pose Estimation 9921
Haolun Li, Fuchen Zheng, Ye Liu, Jian Xiong, Wenhua Zhang, Haidong Hu, Hao Gao

Enhanced Loudspeaker Membrane Excursion Control Method Using Low Latency Distortion Prediction and Efficient LSTM Networks 9926
Rishabh Gupta, Sandeep Agri, Yughendaran Palanivel, Omsrinath Chelamkuri, Raj Narayana Gadde

Optimizing Multimodal Image Fusion: A Novel Approach with Nystrom Attention Mechanisms in Transformer Models 9931
Yuqin Zeng, Ze Wen, Shuqian Fan

Twenty-Five Years of MIR Research: Achievements, Practices, Evaluations, and Future Challenges 9936
Geoffroy Peeters, Zafar Rafii, Magdalena Fuentes, Zhiyao Duan, Emmanouil Benetos, Juhan Nam, Yuki Mitsufuji

Uncertainty-Aware Dynamic Fusion for Multimodal Clinical Prediction Tasks 9941
Jiayu Guo, Ying Cheng, Wen He, Yuejie Zhang, Rui Feng, Xiaobo Zhang

Domain-Incremental Learning for Audio Classification 9946
Manjunath Mulimani, Annamaria Mesaros

TS-SUPERB: A Target Speech Processing Benchmark for Speech Self-Supervised Learning Models 9951
Junyi Peng, Takanori Ashihara, Marc Delcroix, Tsubasa Ochiai, Oldrich Plchot, Shoko Araki, Jan Černocký

Jointly Optimizing Data Discretization and Naive Bayes Classifier Via Multi-Objective Optimization 9956
Jiacheng Tu, Haiyan Yu, Ruxin Ding, Shihe Wang, Jianfeng Ren, Xudong Jiang

FedSe: Group-Based Sequential Training Strategies for Mitigating Label Skew in Federated Learning 9961
Ketu Qiao, Yi Wang, Baoquan Wang, Zhengdong Luo, Xi Zhou

Investigation of Whisper ASR Hallucinations Induced by Non-Speech Audio 9966
Mateusz Barański, Jan Jasiński, Julitta Bartolewska, Stanisław Kacprzak, Marcin Witkowski, Konrad Kowalczyk

Robust Multi-Pitch Estimation Via Optimal Transport Clustering 9971
Anton Björkman, Filip Elvander

CR-CLIP: Image-Text Contrastive Regression for Generalized Gaze Estimation 9976
Yitong Zhu, Xurong Xie, Naiming Yao, Hui Chen, Feng Tian

Mates of Cross Z-Complementary Pairs for Channel Estimation in Generalized SM-MIMO System 9981
Shibsankar Das, Adrish Banerjee

Classification Inconsistency Alignment Network for Cross-Corpus Speech Emotion Recognition 9986
Xiaoyan Zhou, Jiajie Li, Qida Yu, Quan Wu

Chitrarth: Bridging Vision and Language for a Billion People 9991
Shaharukh Khan, Ayush Tarun, Abhinav Ravi, Ali Faraz, Praveen Kumar Pokala, Anagha Bhangare, Raja Kolla, Chandra Khatri, Shubham Agarwal

Dual-Triple Transformer Networks for Accurate CT Pleural Effusion Segmentation 9996
Jianwei Yang, Wenkang Fan, Hao Fang, Zirui Zhu, Xiongbiao Luo

Transformer Based Multi-View Learning for Integrating Static and Dynamic Complementarity of Brain Function 10001
Shengbing Pei, Yan Wang, Zhao Lv, Chao Zhang

Functional Near-Infrared Spectroscopy Feature Extraction with Application in Workload Estimation 10006
Elisabeth R. M. Heremans, David Johnston, Dimitra Emmanouilidou, Andre Golard, Ivan Tashev, Ryen White

Gaze-GZ: Generalized Gaze Estimation with Multi-Scale Gaze Zone Prediction 10011
Zheng Gao, Puneet Kumar, Hao Zou, Xiaobai Li

Sound-Based Recognition of Touch Gestures and Emotions for Enhanced Human-Robot Interaction 10016
Yuanbo Hou, Qiaoqiao Ren, Wenwu Wang, Dick Botteldooren

Leveraging Visual Captions for Enhanced Zero-Shot HOI Detection 10021
Yanqing Zeng, Yunyao Mao, Zhenbo Lu, Wengang Zhou, Houqiang Li

Semantic to Structure: Learning Structural Representations for Infringement Detection 10026
Chuanwei Huang, Zexi Jia, Hongyan Fei, Yeshuang Zhu, Zhiqiang Yuan, Jinchao Zhang, Jie Zhou

Test Time Prompt Tuning for Domain Adaptive Gaze Estimation 10031
Jingjing Wang, Pengwei Yin

Prompt-UIE: A Unified Prompt-Driven Framework for Underwater Image Enhancement 10036
Yanling Zhang, Linxuan Luo, Pan Mu, Cong Bai

Exploring the Implicit Semantic Ability of Multimodal Large Language Models: A Pilot Study on Entity Set Expansion 10041
Hebin Wang, Yangning Li, Yinghui Li, Hai-Tao Zheng, Wenhao Jiang, Hong-Gee Kim

SmartNet: One-Shot Talking Head Synthesis Via Subtle Motion and Appearance Compensation 10046
Wei Hu, Yuzhu Ji, An Zeng, Dan Pan, Yiqun Zhang, Haijun Zhang

Semi-Intrusive Audio Evaluation: Casting Non-Intrusive Assessment as a Multi-Modal Text Prediction Task 10051
Jozef Coldenhoff, Milos Cernak

Low Complexity Riemannian Coordinate-Descent Over Symmetric Positive Definite Matrices 10056
Yogesh Darmwal, Ketan Rajawat

Gen-A: Generalizing Ambisonics Neural Encoding to Unseen Microphone Arrays 10061
Mikko Heikkinen, Archontis Politis, Konstantinos Drossos, Tuomas Virtanen

Near-Field AoA Estimation with Complex Convolutional Kolmogorov-Arnold Network 10066
Jiayi Wang, Disheng Xiao, Yingkai Cao, Kai Ying, Ming Xiao

KGD-GNN: A Knowledge-Guided Graph Neural Network for Myocardial Infarction Localization Via 12-Lead ECG 10071
Lin Guo, Yingqi Wu, Nan Ma, Ying An

Rethinking Encoder-Decoder Flow Through Shared Structures 10076
Frederik Laboyrie, Mehmet Kerim Yucel, Albert Saà-Garriga

Injecting Global Context for Multivariate Time Series Forecasting on Variable Subsets 10081
Xin-Yi Li, Yu-Bin Yang

Projection Modules for Few-Shot Object Detection 10086
Benjamin Rise, Murat Uney, Xiaowei Huang

U2AD: A UAV-Assisted Autonomous Driving Framework for Enhancing Vehicle Risk Perception and Decision-Making Capabilities 10091
Chuangxin Li, Yongqiang Gao, Rao Fu, Jia Chen

Punctuation Restoration: A Case Study of BERT-Based Models' Task-Specific Excellence 10096
Qishuai Zhong, Aixin Sun

CritiPrefill: A Segment-Wise Criticality-Based Approach for Prefilling Acceleration in LLMs..... 10101
Junlin Lv, Yuan Feng, Xike Xie, Xin Jia, Qirong Peng, Guiming Xie

A Lightweight and Real-Time Binaural Speech Enhancement Model with Spatial Cues Preservation 10106
Jingyuan Wang, Jie Zhang, Shihao Chen, Miao Sun

Hierarchical Spatial-Temporal Enhancement Network for Continuous Sign Language Recognition 10111
Min Xu, Sheng Liu, Yuan Feng, Yiheng Yu, Zhelun Jin, Xuhua Yang

Sagalee: An Open Source Automatic Speech Recognition Dataset for Oromo Language 10116
Turi Abu, Ying Shi, Thomas Fang Zheng, Dong Wang

MIFAE-Forensics: Masked Image-Frequency AutoEncoder for DeepFake Detection..... 10121
Hanyi Wang, Zihan Liu, Shilin Wang

Audio-Driven Reinforcement Learning for Head-Orientation in Naturalistic Environments 10126
Wessel Ledder, Yuzhen Qin, Kiki van der Heijden

DiffuseFIST: A Fast Image-Guided Style Transfer Method for Adapting Large-Scale Diffusion Models..... 10131
Miaomiao Dai, Qianyu Zhou, Ran Yi, Lizhuang Ma

Topological Scattering Over Product Cell Complexes 10136
Ayushman Raghuvanshi, Sravanthi Gurugubelli, Sundeep Prabhakar Chepuri

Linguistics-Vision Monotonic Consistent Network for Sign Language Production..... 10141
Xu Wang, Shengeng Tang, Peipei Song, Shuo Wang, Dan Guo, Richang Hong

A GNSS-IR Aided Multispectral Satellite Data Fusion for Meter-Level Wide-Area Volumetric Soil Moisture Estimation 10146
Nicolás Padrón, Sergiy Vorobyov

Adversarial Training and Gradient Optimization for Partially Deepfake Audio Localization..... 10151
Siding Zeng, Jiangyan Yi, Jianhua Tao, Jiayi He, Zheng Lian, Shan Liang, Chuyuan Zhang, Yujie Chen, Xiaohui Zhang

AdaBoost-Based Channel Estimation in One-Bit Millimeter-Wave MIMO 10156
Majdoddin Esfandiari, Petteri Pulkkinen, Sergiy A. Vorobyov, Visa Koivunen

Diversified Augmentation with Domain Adaptation for Debiased Video Temporal Grounding 10161
Junlong Ren, Gangjian Zhang, Haifeng Sun, Hao Wang

CurMIM: Curriculum Masked Image Modeling 10166
Hao Liu, Kun Wang, Yudong Han, Haocong Wang, Yupeng Hu, Chunxiao Wang, Liqiang Nie

Archetypal Analysis for Binary Data..... 10171
A. Emilie J. Wedenborg, Morten Mørup

Deep Diffusion Gradients Leakage in Federated Learning 10176
Dexuan Chen, Yueyi Luo, Qianqian Qi, Hongxiao Fei

Self-Supervised Hyperspectral and Multispectral Fusion Via Deep Low-Rank Prior and Learnable Degradation Networks 10181
Na Liu, Lianming Xu, Suxian Fu, Li Wang

Private Semantic Communications with Separate Blind Encoders 10186
Amirreza Zamani, Mikael Skoglund

A Decade of DCASE: Achievements, Practices, Evaluations and Future Challenges 10191
Annamaria Mesaros, Romain Serizel, Toni Heittola, Tuomas Virtanen, Mark D. Plumbley

Large Multimodal Model is a Better Comparator on Facial Beauty Prediction 10196
Zhenyou Liu, Xuefeng Liang, Jian Lin

Hierarchical Multimodal Decoupling-Fusion Framework for Offline Multiple Appropriate Facial Reaction Generation 10201
Qincheng Lv, Xiaofeng Liu, Jie Li, Rongrong Ni, Pujun Xue, Siyang Song

Multi-Grained Feature Pruning for Video-Based Human Pose Estimation 10206
Zhigang Wang, Shaojing Fan, Zhenguang Liu, Zheqi Wu, Sifan Wu, Yingying Jiao

Problem Solving-Oriented Programming Knowledge Tracing from Behavior to Thought 10211
Tao Huang, Linxia Tang, Huali Yang, Xinjia Ou, Shengze Hu, Jing Geng, Junjie Hu

Sequential Decoding of Multiple Traces Over the Syndrome Trellis for Synchronization Errors 10216
Anisha Banerjee, Lorenz Welter, Alexandre Graell I Amat, Antonia Wachter-Zeh, Eirik Rosnes

Bi-Attention Pyramid Network for Small Defect with Complex Background in Industrial Detection 10221
Yihang Li, Zhiyuan Zou, Xu Liang

An Explicit Consistency-Preserving Loss Function for Phase Reconstruction and Speech Enhancement 10226
Pin-Jui Ku, Chun-Wei Ho, Hao Yen, Sabato Marco Siniscalchi, Chin-Hui Lee

Stacking U-Nets in U-Shape: Redesigning the Information Flow in Model-Based Networks for MRI Reconstruction 10231
Xiaoyu Qiao, Weisheng Li, Bin Xiao, Yuping Huang, Lijian Yang

Dynamic Soft Contrastive Learning for Time Series Anomaly Detection 10236
Yifan Song, Yu Liu, Shaolong Shu

MVDC : A Multi-View Dental Completion Model Based on Contrastive Learning 10241
Xunyu Yang, Qingxin Deng, Minghan Huang, Landu Jiang, Dian Zhang

Edge-Interaction Mamba Network for MRI Brain Tumor Segmentation 10246
Muqing Zhang, Qiule Sun, Yutong Han, Jianxin Zhang

GeMIMO: Searching the Cores of X-Formers for Time Series Forecasting 10251
Zhicheng Zhang, Yong Wang, Shaoqi Tan, Yujie Luo

GBA-Net: A Method for 3D Brain Tumor Segmentation Based on Multi-Scale Gaussian Boundary Attention 10256
Wei Wang, Longrun Wang, Xin Wang

Efficient Extreme Large-Scale Speaker Verification: Dynamic Active Sub Fully-Connected Layers for Faster Training and Memory Optimization 10261
Fulin Zhang, Chenguang Hu, Yao Shen, Yingying Gao, Shilei Zhang, Junlan Feng

Improvements of Discriminative Feature Space Training for Anomalous Sound Detection in Unlabeled Conditions 10266
Takuya Fujimura, Ibuki Kuroyanagi, Tomoki Toda

WMCodec: End-To-End Neural Speech Codec with Deep Watermarking for Authenticity Verification 10271
Junzuo Zhou, Jiangyan Yi, Yong Ren, Jianhua Tao, Tao Wang, Chu Yuan Zhang

Dynamic Structure Hypergraph for Document-Level Event Extraction 10276
Qi Ren, Weihua Wang, Jie Yu, Guanglai Gao

Divide-And-Conquer Variational Bayesian Inference for Multi-Task Learning of High-Resolution SAR Imagery 10281
Lei Yang, Ming Sun, Zhongwei Hu, Zenan Zhang, Wenxuan Yuan

STAR: A Spatial-Temporal Autoencoder for EEG Restoration in Emotion Recognition 10286
Hao-Long Yin, Wei-Long Zheng, Bao-Liang Lu

Frequency Domain Information Integrated Network for Low-Light Image Enhancement 10291
Na Li, Xi Luo, Dunlu Peng, Zied Bouraoui

Learning Two-Factor Representation for Magnetic Resonance Image Super-Resolution 10296
Weifeng Wei, Heng Chen, Pengxiang Su

Probabilistic Contrastive Test-Time Adaptation 10301
Linjing You, Jiabao Lu, Xiayuan Huang

Heterogeneous Graph Convolutional Neural Networks for EEG-FNIRS Bimodal Emotion Recognition 10306
Tao Zhao, Yunlong Xue, Jie Zhu, Wenming Zheng

MTPareto: A MultiModal Targeted Pareto Framework for Fake News Detection 10311
Kaiying Yan, Moyang Liu, Yukun Liu, Ruibo Fu, Zhengqi Wen, Jianhua Tao, Xuefei Liu, Guanjun Li

MIMO Channel as a Neural Function: Implicit Neural Representations for Extreme CSI Compression 10316
Haotian Wu, Maojun Zhang, Yulin Shao, Krystian Mikolajczyk, Deniz Gündüz

Dike: Enhancing Fairness and Efficiency in GPU Clusters for Deep Learning 10321
Bingting Jiang, Jing Yao, Heyi Mu, Xin Su, Zhuo Tang

Massive MIMO-ISAC Beamforming Design Via Sensing Energy Maximization 10326
Fang Li, Bin Liao

FAF-Filt: Frequency-Aware Fourier Filter for Sound Event Detection 10331
Siyu Sun, Xiaohuai Le, Zhuangqi Chen, Xianjun Xia, Chuanzeng Huang

TextureDiffusion: Target Prompt Disentangled Editing for Various Texture Transfer 10336
Zihan Su, Junhao Zhuang, Chun Yuan

Enhancing Incomplete Multimodal Learning Via Modal Complementary Recovering 10341
Meirong Ding, Hongyi Lin, Jiawei Zhu, Chuang Zou, Wenxiu Cai, Bingzhi Chen

Unsupervised Domain Adaptation for Music Transcription: Exploiting Cross-Version Consistency 10346
Lele Liu, Christof Weiß

Leveraging Audio-Only Data for Text-Queried Target Sound Extraction 10351
Kohei Saijo, Janek Ebbers, François G. Germain, Sameer Khurana, Gordon Wiehern, Jonathan Le Roux

Appearance-Adapter: A Self-Supervised Pose-Guided Human Image Synthesis Approach 10356
Sameer Malik, Moyuru Yamada

Implicit and Explicit Rule Injection for Complex Query Answering Over Knowledge Graphs 10361
Xin Zhang, Zhe Wang, Guozheng Rao, Kewen Wang

DPI-TTS: Directional Patch Interaction for Fast-Converging and Style Temporal Modeling in Text-To-Speech 10366
Xin Qi, Ruibo Fu, Zhengqi Wen, Tao Wang, Chunyu Qiang, Jianhua Tao, Chenxing Li, Yi Lu, Shuchen Shi, Zhiyong Wang, Xiaopeng Wang, Yuankun Xie, Yukun Liu, Xuefei Liu, Guanjun Li

Identity-Preserving Diffusion for Face Restoration 10371
Xiaying Bai, Yuhao Yang, Wenming Yang, Rui Zhu, Jing-Hao Xue

Online Learning with Non-Convex Losses: New Condition to Achieve Small Dynamic Regret 10376
Sumit Sah, B. N. Bharath

C3D-VIT: Consistency-Aware 3D Vision Transformer for Face Forgery Detection 10381
Jingyi Zhang, Peng Zhang, Jingjing Wang

Explaining Speaker and Spoof Embeddings Via Probing 10386
Xuechen Liu, Junichi Yamagishi, Md Sahidullah, Tomi Kinnunen

Dual-Space Augmented Intrinsic-LoRA for Wind Turbine Segmentation 10391
Shubh Singhal, Raül Pérez-Gonzalo, Andreas Espersen, Antonio Agudo

Mitigating Hallucinations on Object Attributes Using Multiview Images and Negative Instructions 10396
Zhijie Tan, Yuzhi Li, Shengwei Meng, Xiang Yuan, Weiping Li, Tong Mo, Bingce Wang, Xu Chu

Simplified One-Sided Image-To-Image Translation with Reconstruction-Constrained Generative Adversarial Networks 10401
Shuocheng Wang, Qingfeng Wu, Mengyuan Ge, Yingdong Wang

Multilingual Parameter-Sharing Adapters: A Method for Optimizing Low-Resource Neural Machine Translation 10406
Yunlong Zhang, Nan Chen, Yonghe Wang, Xiangdong Su, Feilong Bao

SSAST-Adapter: A Parameter-Efficient Incremental Learning Algorithm for Underwater Acoustic Target Recognition 10411
Qian Zhu, Qisheng Xu, Boqing Zhu, Zijian Gao, Lingbin Zeng, Kele Xu

Robust Low-Light Human Pose Estimation Through Illumination-Texture Modulation 10416
Feng Zhang, Ze Li, Xiatian Zhu, Lei Chen

Pan-Protein Design Learning Enables Task-Adaptive Generalization for Low-Resource Enzyme Design 10421
Jiangbin Zheng, Ge Wang, Han Zhang, Stan Z. Li

CorrGAN: Simultaneous Learning of Speech Enhancement and Perceptual Quality Loss Functions 10426
Vasily Zadorozhnyy, Saeed Amizadeh, Qiang Ye, Kazuhito Koishida

DDA: Distillation-Driven Acceleration of the Reverse Diffusion Process for Stochastic Multi-Ship Trajectory Prediction 10431
Kun Ma, Qilong Han, Jingzheng Yao, Changmao Wu, Yuntao Zhang

BDGAN: Boundary and Diversity-Aware Generative Adversarial Network for Imbalanced Medical Image Augmentation 10436
Hongwei Ding, Qi Tao, Nana Huang

Dark Experience for Incremental Keyword Spotting 10441
Tianyi Peng, Yang Xiao

Fast Structured Orthogonal Dictionary Learning Using Householder Reflections..... 10446
Anirudh Dash, Aditya Siripuram

Online Waveform Selection for Cognitive Radar..... 10451
Thulasi Tholeti, Avinash Rangarajan, Sheetal Kalyani

Sparse PCA with Oracle Rate in High Dimensions..... 10456
Wenfu Zhong, Ziping Zhao

Noise-Resilient Unlimited Sampling and Recovery of Sparse Signals 10461
Geethu Joseph

Boosting Text-To-Image Generation Via Multilingual Prompting in Large Multimodal Models 10466
Yongyu Mu, Hengyu Li, Junxin Wang, Xiaoxuan Zhou, Chenglong Wang, Yingfeng Luo, Qiaozhi He, Tong Xiao, Guocheng Chen, Jingbo Zhu

EFL-PEFT: A Communication Efficient Federated Learning Framework Using PEFT Sparsification for ASR..... 10471
Mohamed Nabih Ali, Daniele Falavigna, Alessio Brutti

V2SFlow: Video-To-Speech Generation with Speech Decomposition and Rectified Flow 10476
Jeongsoo Choi, Ji-Hoon Kim, Jinyu Li, Joon Son Chung, Shujie Liu

Learning Time-Varying Graphs from Data with Few Causes..... 10481
Panagiotis Misiakos, Markus Püschel

Disentangling Hierarchical Features for Anomalous Sound Detection Under Domain Shift..... 10486
Jian Guan, Jiantong Tian, Qiaoxi Zhu, Feiyang Xiao, Hejing Zhang, Xubo Liu

ArtTwin: A Novel Concept of Developing Digital Twin of Human Arterial System 10491
Akasapu Hemanthika, Dinu S Chandran, Sandeep Kumar, Sitikantha Roy

Offline Reinforcement Learning Via Conservative Smoothing and Dynamics Controlling..... 10496
Haihong Guo, Fengxin Li, Jiao Li, Hongyan Liu

Sufficient Learning for Label Noise with Dual-Regularization..... 10501
Yongfeng Dong, Guifang Wu, Jiaji Wang, Zhen Wang

VisAgent: Narrative-Preserving Story Visualization Framework..... 10506
Seungkwon Kim, GyuTae Park, Sangyeon Kim, Seung-Hun Nam

A Blind Super-Resolution Method for Near-Field Channel Estimation with Angle-Range Recovery 10511
Jie Li, Qiaozhen Chen, Feng Xi

MDDNet: Multilevel Difference-Enhanced Denoise Network for Unsupervised Change Detection in SAR Images 10516
He Zong, Erlei Zhang, Xinyu Li, Hongming Zhang, Jinchang Ren

Tracking Time-Varying Parameters in Massive MIMO IoT Networks: A Linear Coherent Decentralized Approach 10521
Kunwar Pritiraj Rajput, Linlong Wu, Bhavani Shankar M. R., Bjorn Ottersten, Pramod K. Varshney

PiCNet: Physics-Infused Convolution Network for Radar-Based Precipitation Nowcasting 10526
Zheng Wang, Hanyi Zhang, Cong Bai

DSDN-Net: An Effective Network for Semantic Segmentation in Open-Pit Coal Mining Areas for Land Cover Recognition 10531
Jiaqi Li, Ming Ma

LABEL-SAM: A Semi-Automatic Interactive Annotation Model for Aortic Dissection Segmentation in 3D CTA Image 10536
Wenjie Cai, Tao Tang, Balachander.J, Lingming Kong, Ying Zhou, Qingfeng Wang, Jing Li

ITMO Language Diarization and Identification Systems for the DISPLACE 2024 Challenge 10541
Egor Ausev, Vladimir Volokhov, Sergey Novoselov, Vladislav Marchevskiy, Ekaterina Shangina, Alexey Logunov

A Prompt Learning Framework with Large Language Model Augmentation for Few-Shot Multi-Label Intent Detection 10546
Ning Zhuang, Xiao Wei, Junlei Li, Xiaobao Wang, Chenyang Wang, Longbiao Wang, Jianwu Dang

Poisoning the Diffusion: A Simple and Robust Watermarking Method for Audio Generation 10551
Yi Tang

Adapting Large Language Models to Forecast in Frequency Domain 10556
Yungeng Zhang, Yuan Chang, Xiaohou Shi, Yaqi Song, Feng Wang, Mingchuan Yang

LMFCA-Net: A Lightweight Model for Multi-Channel Speech Enhancement with Efficient Narrow-Band and Cross-Band Attention 10561
Yaokai Zhang, Hanchen Pei, Wanqi Wang, Gongping Huang

FKAN-GMFNet: Fourier Kolmogorov-Arnold-Based Group Multi-Scale Fusion Network for Aneurysm Image Segmentation 10566
Shanchen Pang, Xue Zhao, Yulin Zhang, Yawu Zhao, Hengtao Ding, Zhiyuan Zhao, Sibo Qiao

Deepfake Detection of Singing Voices with Whisper Encodings 10571
Falguni Sharma, Priyanka Gupta

BiCG: Binaural Cue Generation from Unified HRTF Datasets 10576
Xikun Lu, Yilei Wang, Jinqiu Sang, Chengshi Zheng

Efficient Modeling and Low Complexity Implementation of Rate Estimation in Versatile Video Coding 10581
Yanze Wang, Tianyi Sun, Hui Zhao, Jun Sun

APLASE: Compression Using Adaptive Piecewise Linear Approximation and Sparse Encoding 10586
Ravi Raj Saxena, T.V. Prabhakar, Joy Kuri, Chandra R. Murthy

Harnessing the Zero-Shot Power of Instruction-Tuned Large Language Model for Guiding End-To-End Speech Recognition 10591
Yosuke Higuchi, Tetsuji Ogawa, Tetsunori Kobayashi

Diffusion Counterfactual-Based Anomaly Detection in Class-Imbalanced Data 10596
Xinyun Shen, Min Li, Zhengmao Ye, Zhenyang Yu, Lei Duan

GMM-Based Bootstrap Prototype-Aware Learning for Weakly Supervised Semantic Segmentation 10601
Yahui Wu

Enhancing Change Detection in Remote Sensing: Integrating Synthetic Data with Semi-Supervised Learning 10606
Yafei Luo, Erik Meijering, Yang Song

AUIED3K: A New Andaman Underwater Image Enhancement Dataset for Deep Learning-Driven Image Enhancement with Minimum Loss Dehazing 10611
Praveen Saini, Navjot Singh, Anshu S. Anand

NeRF-3DTalker: Neural Radiance Field with 3D Prior Aided Audio Disentanglement for Talking Head Synthesis 10616
Xiaoxing Liu, Zhilei Liu, Chongke Bi

WinStega: An Adaptive Robust Enhancement Framework for Generative Linguistic Steganography 10621
Kaiyi Pang, Minhao Bai, Jinshuai Yang, Wei-Qiang Zhang, Minghu Jiang, Yongfeng Huang

Layer-Animate for Transparent Video Generation 10626
Jingqi Bai, Jingkai Zhou, Benzhi Wang, Weihua Chen, Yang Yang, Zhen Lei, Fan Wang

MA-Det: A Discriminative Morphology-Aware Detector for Cervical Lesion Cell Clumps 10631
Ziyang Yin, Qian Huang, Yulin Chen, Hao Lu

Doppler Single-Photon Lidar 10636
Ruangrawee Kitichotkul, Joshua Rapp, Yanting Ma, Hassan Mansour

Two-In-One: Unified Multi-Person Interactive Motion Generation by Latent Diffusion Transformer 10641
Boyuan Li, Xihua Wang, Ruihua Song, Wenbing Huang

Advancing Non-Intrusive Suppression on Enhancement Distortion for Noise Robust ASR 10646
Wei Wang, Siyi Zhao, Yanmin Qian

MalImgDA: Diffusion-Based Data Augmentation for Long-Tailed Malware Family Classification 10651
Gang Yang, Jun He, Bo Wu, Tao Xia, Linna Fan, Lin Ni

Latent Space Score-Based Diffusion Model for Probabilistic Multivariate Time Series Imputation 10656
Guojun Liang, Najmeh Abiri, Atiye Sadat Hashemi, Jens Lundström, Stefan Byttner, Prayag Tiwari

Structural Similarity-Aware Cross-Domain Transformer for Improved Seismic Fault Detection 10661
Tiash Ghosh, Razeen A. Rasheed, Sanjai Kumar Singh, Mamata Jenamani, Aurobinda Routray

Extract Information from Hybrid Long Documents Leveraging LLMs: A Framework and Dataset 10666
Chongjian Yue, Xinrun Xu, Xiaojun Ma, Lun Du, Zhiming Ding, Shi Han, Dongmei Zhang, Qi Zhang

Prototype-Driven Multi-Feature Generation for Visible-Infrared Person Re-Identification 10671
Jiarui Li, Qiu Zhen, Yilin Yang, Yuqi Li, Zeyu Dong, Chuanguang Yang

MixHD: A Method for Detecting Hallucinations Based on the Internal State and Output Probability of Large Language Models 10676
Chuang Li, Bingnan Xing, Dongdong Huo, Qihui Zhou, Zhen Xu, Yu Wang

FBSE-FTFCWT-Based Novel Automated Framework for Dysarthric Speech Detection 10681
Amishi Vijay, Ram Bilas Pachori, Balasubramanyam Appina, Nitya Tiwari

How Redundant is the Transformer Stack in Speech Representation Models? 10686
Teresa Dorszewski, Albert Kjøller Jacobsen, Lenka Tětková, Lars Kai Hansen

Geometry-Constrained EEG Channel Selection for Brain-Assisted Speech Enhancement 10691
Keying Zuo, Qingtian Xu, Jie Zhang, Zhenhua Ling

A Novel Single Continuous Shot Multiple Lesions Endoscopy Report Generation 10696
Xinpan Yuan, Junhua Kuang, Liujie Hua, Guihu Zhao, Changhong Zhang, Jiabao Li

ST-HCSS: Deep Spatio-Temporal Hypergraph Convolutional Neural Network for Soft Sensing 10701
Hwa Hui Tew, Fan Ding, Gaoxuan Li, Junn Yong Loo, Chee-Ming Ting, Ze Yang Ding, Chee Pin Tan

Enhancing EEG-Based Covert Speech Decoding Through Knowledge Transfer 10706
Zhiwei Guo, Muyun Jiang, Chenyu Liu, Min Wu, Jia Lu, Balázs Gulyás, Cuntai Guan

MS-UFAD: A Large-Scale Dataset for Real-World Unified Face Attack Detection with Text Descriptions 10711
Ning Jiang, Dingheng Zeng, Liang Gao, Sheng Chen, Zhifei Kong, Yanhong Liu, Jiao Li, Yue Feng, Tongtong Yuan, Weihong Deng, Quan Lu, Ying Li

Dynamic Category Queries Transformer for Generalized Few-Shot Semantic Segmentation 10716
Kunze Huang, Jieyuan Yang, Andreas Jakobsson, Luyao Tang, Xiaotong Tu, Xinghao Ding, Yue Huang

Collaborative Dual-Branch Spatial-Frequency Enhancement Network for Low-Light Images 10721
Tao He, Tiecheng Song, Yin Liu, Feng Yang, Ruiyuan Chen, Zhixin Li

Enhancing Complex Formula Recognition with Hierarchical Detail-Focused Network 10726
Jiale Wang, Junhui Yu, Huanyong Liu, Chenanran Kong

Accurate Hardware Trojan Detection for SGIN Device: A Prompt-Tuning and LangChain Approach 10731
Duo Zhang, Ming Mao, Yunchuan Guo, Fenghua Li, Daiyong Quan

Certainty-Guided Reasoning and Refinement Network for Camouflaged Object Detection 10736
Bifan Lai, Meijun Sun, Junkun Zhao, Yan Zhou

A Geometry-Based Node Activation Method for Relative Localization 10741
Licheng Wang, Yi Li, Hanying Zhao, Yuan Shen

Dual Multi-Scale GCN with Deformable Temporal Kernel for Skeleton-Based Action Recognition 10746
Jianan Li, Yangtao Zhou, Hua Chu, Han Wang, Zhifu Zhao, Fei Li, Qingshan Li

Translating Mental Imaginations into Characters with Codebooks and Dynamics-Enhanced Decoding 10751
Jingyuan Li, Yansen Wang, Nie Lin, Dongsheng Li

OF-AR Relation Aware Representation Learning for Lesion Image Segmentation and Grading 10756
Xinpan Yuan, Siming Jin, Liujie Hua, Guihu Zhao, Changhong Zhang, Yuan Guo

Graph Embedded Stochastic Configuration Networks for Imbalanced Data Classification..... 10761
Yuanhang Qiu, Dianhui Wang

DapPep: Domain Adaptive Peptide-Agnostic Learning for Universal T-Cell Receptor-Antigen Binding Affinity Prediction 10765
Jiangbin Zheng, Qianhui Xu, Ruichen Xia, Stan Z. Li

The Conformer Encoder May Reverse the Time Dimension..... 10770
Robin Schmitt, Albert Zeyer, Mohammad Zeineldeen, Ralf Schlűter, Hermann Ney

Capturing Rich Behavior Representations: A Dynamic Action Semantic-Aware Graph Transformer for Video Captioning 10775
Caihua Liu, Xu Li, Wenjing Xue, Wei Tang, Xia Feng

Learning Strategy with Barlow Twins Objective for Emotion-Robust Speaker Verification System..... 10780
Dmitrii V. Mikhailovskii, Aki Kunikoshi, David A. van Leeuwen, Jaebok Kim

Evaluating the Security of Public Surrogate Watermark Detectors 10785
Chloé Imadache, Eva Giboulot, Teddy Furon

Enhancing Cross-Domain Slot Filling with Joint LLM Data Generation and Data Curation 10790
Peijie Huang, Weizhen Li, Yuhong Xu, Junbao Huang

GNCL: A Graph Neural Network with Consistency Loss for Segment-Level Spoofed Speech Detection 10795
Zirui Ge, Xinzhou Xu, Haiyan Guo, Zhen Yang, Björn Schuller

Boosting Lightweight Camouflaged Object Detection with Multi-Scale Context and Boundary Awareness..... 10800
Zihan Xu, Zheng Wang, Haoyu Wang, Cheng Liu, Yan Zhou, Meijun Sun

Consensus Graph-Based Spectral Ensemble Clustering Via Low-Rank Tensor Learning 10805
Zhe Cao, Haonan Xin, Zihua Zhao, Jie Wang, Rong Wang

DEP-SLAM: A Dynamic Environment Perception SLAM System with Large Language Models..... 10810
Ying He, E Richard Yu, Fei Ma, Ming Li, Guang Zhou

PD-SDF: Dynamic Surface Reconstruction Based on Plane Decomposition for Single View RGB-D Videos..... 10815
Jinwen Li, Weixing Xie, Junfeng Yao, Shaoqi Wu, Youhong Peng, Mengyuan Ge, Xiao Dong

KLFormer: Karhunen-Loève Transform for Robust 3D Human Pose Estimation 10820
Xin Zeng, Haonan Luo, Zihang Wang, Sijia Li, Leyu Zhang, Tianrui Li

Principal Curvatures Estimation with Applications to Single Cell Data 10825
Yanlei Zhang, Lydia Mezrag, Xingzhi Sun, Charles Xu, Kincaid Macdonald, Dhananjay Bhaskar, Smita Krishnaswamy, Guy Wolf, Bastian Rieck

Comprehensive Perturbation Consistency for Semi-Supervised Change Detection in Remote Sensing Images..... 10830
Zan Mao, Xin Li, Ze Luo, Yingjuan Tang, Dongmei Jiang

LoVA: Long-Form Video-To-Audio Generation 10835
Xin Cheng, Xihua Wang, Yihan Wu, Yuyue Wang, Ruihua Song

Adaptive Large Language Models Via Attention Shortcuts 10840
Prateek Verma, Mert Pilanci

FARE: A Deep Learning-Based Framework for Radar-Based Face Recognition and Out-Of-Distribution Detection 10845
Sabri Mustafa Kahya, Boran Hamdi Sivrikaya, Muhammet Sami Yavuz, Eckehard Steinbach

Prototypical Part Transformer for Interpretable Image Recognition 10850
Anni Yu, Yu-Bin Yang

Self-Supervised Graph Representation Learning for In-The-Wild Wearable and Smartphone Based Emotion Recognition 10855
Ioannis Ziogas, Leontios J. Hadjileontiadis, Ahsan H. Khandoker, Aamna Al Shehhi

FGU3R: Fine-Grained Fusion Via Unified 3D Representation for Multimodal 3D Object Detection 10860
Guoxin Zhang, Ziying Song, Lin Liu, Zhonghong Ou

Generalization Guarantee of Decentralized Learning with Heterogeneous Data 10865
Haoxiang Ye, Tao Sun, Qing Ling

U-SAM: Upgrade Segment Anything Model with Semantic-Aware and Memory-Efficient 10870
Xiaofeng Jin, Jie Hu, Jianghang Lin, Shengchuan Zhang, Liujuan Cao

Clutter Resilient Occlusion Avoidance for Tightly-Coupled Motion-Assisted Detection 10875
Zhixuan Xie, Jianjun Chen, Guoliang Li, Shuai Wang, Kejiang Ye, Yonina C. Eldar, Chengzhong Xu

Multi-Task Model Fusion Via Adaptive Merging 10880
Luming Chen, Ziwei Xiang, Kai Lei, Xu-Yao Zhang

Deep Dynamic Probabilistic Canonical Correlation Analysis 10885
Shiqin Tang, Shujian Yu, Yining Dong, S. Joe Qin

Guiding Inter-Domain Class Balancing with Salient Features for Domain Adaptive Object Detection 10890
Haiming Peng, Dingkang Yang, Mingxu Wang, Weilong Lin, Xinhua Zeng

KANGAN-AVSS: Kolmogorov-Arnold Network Based Generative Adversarial Networks for Audio-Visual Speech Synthesis 10895
Subhayu Ghosh, Swapnil Saha, Nanda Dulal Jana

Contrast-Unity for Partially-Supervised Temporal Sentence Grounding 10900
Haicheng Wang, Chen Ju, Weixiong Lin, Chaofan Ma, Shuai Xiao, Ya Zhang, Yanfeng Wang

A Deformable-Based Source-Free Unsupervised Domain Adaptation Method for Cervical Cell Detection 10905
Qiao Pan, Yawen Xue, Bin Yang

Standardization Status of MPEG Video-Based Dynamic Mesh Coding (V-DMC) 10910
Wenjie Zou, Shizhuo Zhang, Fuzheng Yang, Marius Preda

A Self-Supervised UAV Detection Method Based on Channel State Information 10915
Pengxuan Gao, Disheng Xiao, Ruiheng Zou, Kai Ying

Situation-Aware Space-Time Waveform Design for Automotive MIMO Radars 10920
Edoardo Focante, Nitin Jonathan Myers, Geethu Joseph, Ashish Pandharipande

Explore the Hallucination on Low-Level Perception for MLLMs 10925
Yinan Sun, Zicheng Zhang, Haoning Wu, Xiaohong Liu, Weisi Lin, Guangtao Zhai, Xiongkuo Min

Delayed Fusion: Integrating Large Language Models into First-Pass Decoding in End-To-End Speech Recognition 10930
Takaaki Hori, Martin Kocour, Adnan Haider, Erik McDermott, Xiaodan Zhuang

Enhancing Large Language Model Inference Efficiency Via Lookahead Cache Filtering 10935
Jie Ou, Yueming Chen, Shuaihong Jiang, Wenhong Tian

Joint Edge and Regional Depth Enhancement Network for Camouflaged Object Detection 10940
Miao Qi, Zheng Wang, Cheng Liu, Yan Zhou, Meijun Sun

Prototypical Graph Alignment for Text-Based Person Search 10945
Yu Huang, Canlong Zhang, Zhixin Li, Zhiwen Wang, Chunrong Wei

Task-Aware Unified Source Separation 10950
Kohei Saijo, Janek Ebbers, François G. Germain, Gordon Wichern, Jonathan Le Roux

A Unified Hardware Accelerator for Fast Fourier Transform and Number Theoretic Transform 10955
Rishabh Shrivastava, Chaitanya Prasad Ratnala, Durga Manasa Puli, Utsav Banerjee

Pathological Section Staining Transferring with Tailored Metric-Based Model Selection 10960
Yiming Ji, Suyang Zhu, Dong Zhang, Shoushan Li

DeepPEM-AFC: An Improved Prediction-Error-Method-Based Adaptive Feedback Cancellation with Deep Learning for Hearing Aids 10965
Xiaofan Zhan, Fengyuan Hao, Xiaodong Li, Chengshi Zheng

Multi-Period Normalization for Long-Term Time Series Forecasting 10970
Jiayu Zhang, Yuantong Dong

Whisper-GPT: A Hybrid Generative LLM for Speech and Music 10975
Prateek Verma

Dimensionality-Reduced Spatial Bipartite Graph Clustering for Hyperspectral and LiDAR Data 10980
Zhe Cao, Haonan Xin, Bo Yan, Jinping Sui, Rong Wang

Reconstruction of EEG and ECG from Single Channel Mixture Using Branched Autoencoder Based Separable Representations 10985
Shreyasi Datta, Jayavardhana Gubbi, Arpan Pal

DiffRS: An Extensible Diffusion Model for Remote Sensing Image Generation 10990
Xinyue Huang, Xin Niu, Jingfei Jiang, Hengyue Pan

Efficient Multi-Branch Black-Box Semantic-Aware Targeted Attack Against Deep Hashing Retrieval 10995
Chihan Huang, Xiaobo Shen

RelaI2P: Relational Learning for Image-To-Point Cloud Registration 11000
Minghui Hou, Gang Wang, Zhiyang Wang, Baorui Ma

XDGesture: An xLSTM-Based Diffusion Model for Co-Speech Gesture Generation 11005
Zixing Zhang, Jiajun Li, Bin Wang, Yiming Liu, Huan Zhao, Björn W. Schuller

Co-Training with Progressive Distribution Alignment and Uncertainty-Interactive Relabeling for Semi-Supervised Domain Adaptive Semantic Segmentation 11010
Ruiguo Yu, Yida Wang, Xuewei Li, Xuzhou Fu, Zijian Zhang, Yuan Tian, Jie Gao

A Small-Footprint Acoustic Echo Cancellation Solution for Mobile Full-Duplex Speech Interactions 11015
Yiheng Jiang, Biao Tian

OTFS for Automotive Radars: Waveform Optimization and Ambiguity Function Analysis 11020
Nazila Karimian-Sichani, Mohammad Alaee-Kerahroodi, Maria S. Greco, Fulvio Gini, Bhavani Shankar M. R

A Continual Learning Approach for Embodied Question Answering with Generative Adversarial Imitation Learning 11025
Xin Zeng, Haonan Luo, Zihang Wang, Sijia Li, Zhixuan Shen, Tianrui Li

Popularity and Interest Signal Detection for Sequential Recommendation Denoising 11030
Xuewei Li, Kunyi Yang, Yue Zhao, Tianyi Xu, Jian Yu, Mei Yu, Mankun Zhao

PicoAudio: Enabling Precise Temporal Controllability in Text-To-Audio Generation 11035
Zeyu Xie, Xuenan Xu, Zhizheng Wu, Mengyue Wu

BadRefSR: Backdoor Attacks Against Reference-Based Image Super Resolution 11040
Xue Yang, Tao Chen, Lei Guo, Wenbo Jiang, Ji Guo, Yongming Li, Jiaming He

Pyramid Attention Enhancement Network for Nighttime UAV Tracking 11045
Xiaomin Huang, Zhenhua Wu, Ying Li, Changjing Shang, Qiang Shen

Language Models Can See Better: Visual Contrastive Decoding for LLM Multimodal Reasoning 11050
Yuqi Pang, Bowen Yang, Haoqin Tu, Yun Cao, Zeyu Zhang

Grouped Knowledge Distillation with Adaptive Logit Softening for Speaker Recognition 11055
Chong-Xin Gan, Youzhi Tu, Zezhong Jin, Man-Wai Mak, Kong Aik Lee

Adaptive Decoding for Efficient Automatic Speech Recognition 11060
Xiangnan Ma, Peizhuo Liu, Yuhao Zhang, Kaiqi Kou, Chenghao Gao, Tong Xiao, Jingbo Zhu

Exact Rotation Invariant Robust PCA 11065
Tanmoy Jana, Nikhil Raghav, Angshul Majumdar, Md Sahidullah

AgentPose: Progressive Distribution Alignment Via Feature Agent for Human Pose Distillation 11070
Feng Zhang, Jinwei Liu, Xiatian Zhu, Lei Chen

AudioTime: A Temporally-Aligned Audio-Text Benchmark Dataset 11075
Zeyu Xie, Xuenan Xu, Zhizheng Wu, Mengyue Wu

Self-Tuning Spectral Clustering for Speaker Diarization 11080
Nikhil Raghav, Avisek Gupta, Md Sahidullah, Swagatam Das

General Dynamic Regularization Federated Learning with Hybrid Sharpness-Aware Minimization 11085
Fengchun Zhang, Dongfen Li, Jinshan Lai, Yang Zhang, Fengli Zhang, Ruijin Wang

Analysis and Calibration of Nonlinear Power Amplifiers in Wideband OFDM-Based LEO Satellite Communication System 11090
Kai Ying, Linshan Zhao, Pengcheng Jia, Ming Zhou, Kai Kang

Point Clean-Label Backdoor Attack for Specific Classes Via Feature Entanglement 11095
Yang Wang, Wei Li, Shengbo Chen, Hong Rao, Azman Mohammad

Compressing a Flow-Based Privacy Protection Model Via a Novel Joint Distilling and Pruning Method 11100
Sissi Xiaoxiao Wu, Zhicong Liang, Zehong Huang

Multi Queue for Unsupervised Person Re-Identification 11105
Zhenyuan Lin, Shengyong Xie, Danhua Liu, Weikun Li, Ang Gao, Yubo Dong

Model-Based Machine Learning for Max-Min Fairness Beamforming Design in JCAS Systems 11110
Mengyuan Ma, Tianyu Fang, Nir Shlezinger, A. L. Swindlehurst, Markku Juntti, Nhan Nguyen

ForensiCam-215K: A Large Scale Image and Video Dataset for Forensic Analysis 11115
Suwen Du, Pengpeng Yang, Daniele Baracchi, Jinglian Jin, Dasara Shullani, Alessandro Piva

Analyzing and Reducing Catastrophic Forgetting in Parameter Efficient Tuning 11120
Xinlong Li, Weijieying Ren, Wei Qin, Lei Wang, Tianxiang Zhao, Richang Hong

Large-Scale Recurrent Neural Networks with Fully Homomorphic Encryption for Privacy-Enhanced Speaker Identification 11125
Vele Tosevski, Glenn Gulak

Class Relevance Learning for Out-Of-Distribution Detection 11130
Liguang Zhou, Butian Xiong, Tin Lun Lam, Yangsheng Xu

AceParse: A Comprehensive Dataset with Diverse Structured Texts for Academic Literature Parsing 11135
Huawei Ji, Cheng Deng, Bo Xue, Zhouyang Jin, Jiaxin Ding, Xiaoying Gan, Luoyi Fu, Xinbing Wang, Chenghu Zhou

Typical Vs. Atypical Disfluency Classification: Introducing the IIITH-TISA Corpus and Temporal Context-Based Feature Representations 11140
Priyanka Kommagouni, Vamshiraghusimha Narasinga, Purva Barche, Sai Akarsh C, Anil Vuppala

Attention-Enhanced Feature Fusion Network for No-Reference Image Quality Assessment 11145
Jiliang Ma, Yihua Chen, Pengsheng Huang, Zhenjun Tang

Speech Synthesis Along Perceptual Voice Quality Dimensions 11150
Frederik Rautenberg, Michael Kuhlmann, Fritz Seebauer, Jana Wiechmann, Petra Wagner, Reinhold Haeb-Umbach

River-GEM: Generating and Enhancing Muddy Water Images 11155
Alik Pramanick, Chivukula Sairam Satwik, Sonal Kumar, Akshay Daydar, Arijit Sur

Improved Motion Plane Adaptive 360-Degree Video Compression Using Affine Motion Models 11160
Marina Ritthaler, Andy Regensky, André Kaup

Evaluating Snippet Significance: A Framework for Audio and Text-Based Dialogue Summarization 11165
Anderson de Lima Luiz, Raviteja Boddu, Munir Georges

Communication-Efficient Verifiable and Oblivious Aggregation with Client Dropouts 11170
Zhangshuang Guan, Yulin Zhao, Longyun Yang, Zhiguo Wan, Jinsong Han

SPTU-Lite: An Efficient Auxiliary Diagnostic Approach Combining Spiking Neural Networks and Large Kernel Extractors for ECG Signals 11175
Chuxuan Shan, Xiaohua Wang, Hang Qi, Qingxu Meng, Jiabao Wang, Weijiang Wang, Yueting Shi

Sparse Bayesian Integrated CNN Framework for Enhanced Acoustic Source Localization 11180
Priyadarshini Dwivedi, Gyanajyoti Routray, Rajesh M. Hegde

Privacy-Aware Federated Fine-Tuning of Large Pretrained Models with Just Forward Propagation 11185
Ke Xing, Yanjie Dong, Xiping Hu, Victor C. M. Leung, M. Jamal Deen, Song Guo

Less is Enough: Relation Graph Guided Few-Shot Learning for Multi-Label Aspect Category Detection 11190
Shiman Zhao, Wei Chen, Tengjiao Wang, Jiahui Yao, Dawei Lu, Jiabin Zheng

DeepPreNet: A Deep Learning Pre-Processing Method for Speech Distortion Correction in Parametric Array Loudspeaker 11195
Wenyao Ma, Yunxi Zhu, Fengyuan Hao, Liwen Qin, Fengyi Fan, Jun Yang

Multimodal Machine Learning Can Predict Videoconference Fluidity and Enjoyment 11200
Andrew Chang, Viswadruth Akkaraju, Ray McFadden Cogliano, David Poeppel, Dustin Freeman

Speech Enhancement with Overlapped-Frame Information Fusion and Causal Self-Attention 11205
Yuewei Zhang, Huanbin Zou, Jie Zhu

Appearance- And Orientation-Aware Fine-Grained Rotated Ship Detection in High-Resolution Satellite Imagery 11210
Yan Li, Lingyi Liu, Yunpeng Bai, Ying Li, Qiang Shen

MTE: Multi Transformation of Entities in Quaternion Vector Space for Temporal Knowledge Graph Completion 11215
Ruiguo Yu, Jiazheng Guo, Mankun Zhao, Tianyi Xu, Jiujiang Guo, Wenbin Zhang, Mei Yu

BRDIA: Bidirectional Reasoning with Dynamic Instruction Adjustment for Multi-Hop KGQA 11220
Chuanyang Gong, Zhihua Wei

Dual-Function Waveform Design in Wireless Sensor Networks Via SoS Optimization 11225
Robin Amar, Linlong Wu, Saeid Sedighi, Mohammad Alaee-Keerahroodi, M. R. Bhavani Shankar

Cross-Modal Gaussian Localization Distillation for Optical Information Guided SAR Object Detection 11230
Chao Wang, Lei Luo, Wenxuan Fang, Jian Yang

Enhancing Teacher Classroom Behavior Descriptions: A Spatio-Temporal Graph-Based Method for Video Captioning 11235
Ting Cai, ChengYang He, Yu Xiong

SwinGAN-AVSS: Audio-Visual Speech Synthesis Leveraging Swin Transformer–Enhanced Generative Adversarial Networks 11240
Subhayu Ghosh, Swapnil Saha, Nanda Dulal Jana

Mobile-Friendly Image De-Noising: Hardware Conscious Optimization for Edge Application 11245
Srinivas Miriyala, Sowmya Vajrala, Hitesh Kumar, Sravanth Kodavanti, Vikram Rajendiran

Improving Knowledge Distillation Via Cross-Modal Insights from CLIP 11250
Jingtao Zhou, Hao Zheng, Wenkai Zhong, Zhiqiang Bao

Multi-View Hypergraph-Based Contrastive Learning Model for Cold-Start Micro-Video Recommendation 11255
Sisuo Lyu, Xiuze Zhou, Xuming Hu

Temporal-Frequency State Space Duality: An Efficient Paradigm for Speech Emotion Recognition 11260
Jiaqi Zhao, Fei Wang, Kun Li, Yanyan Wei, Shengeng Tang, Shu Zhao, Xiao Sun

NanoVoice: Efficient Speaker-Adaptive Text-To-Speech for Multiple Speakers 11265
Nohil Park, Heeseung Kim, Che Hyun Lee, Jooyoung Choi, Jiheum Yeom, Sungroh Yoon

LP-Gaussians: Learnable Parametric Gaussian Splatting for Efficient Dynamic Reconstruction of Single-View Scenes 11270
Shaoqi Wu, Weixing Xie, Youhong Peng, Jinwen Li, Jiawei Yao, Junfeng Yao

PeT-KeyStAtion: Parameter-Efficient Transformer with Keypoint-Guided Spatial-Temporal Aggregation for Video-Based Person Re-Identification 11275
Xingan Ma, Jinhui Yi, Juergen Gall

Spatial-Temporal Reconstruction Error for AIGC-Based Forgery Image Detection 11280
Chengji Shen, Zhenjiang Liu, Kaixuan Chen, Jie Lei, Mingli Song, Zunlei Feng

Sample-Level Self-Paced Learning to Tackle Multimodal Imbalance Problem.....11285
Ying Zhou, Xuefeng Liang, Yue Xu, Bowen Gao

FBI-Net: Frequency Band Integration Network for Infrared Small Target Segmentation.....11290
Biqiao Xin, Qiang Li, Qianchen Mao, Jinbao Wang, Bingshu Wang

Classification Error Bound for Low Bayes Error Conditions in Machine Learning 11295
Zijian Yang, Vahe Eminyan, Ralf Schlüter, Hermann Ney

Enabling Auditory Large Language Models for Automatic Speech Quality Evaluation 11300
Siyin Wang, Wenyi Yu, Yudong Yang, Changli Tang, Yixuan Li, Jimin Zhuang, Xianzhao Chen, Xiaohai Tian, Jun Zhang, Guangzhi Sun, Lu Lu, Chao Zhang

Hybrid Predictive and Parametric Stereo Coding for Voice and Audio Communications.....11305
Guillaume Fuchs, Emmanuel Ravelli, Franz Reutelhuber, Eleni Fotopoulou

SACR: Self-Training with Saliency-Augmented Consistency Regularization for Few-Shot Learners.....11310
Yanyan Feng, Yue Zhou, Yun Xue, Fenghuan Li, Zehong Lin

Exploring Meta Evidence for Prompt Optimization.....11315
Yao He, Jiajun Liu, Wenjun Ke, Peng Wang, Congda Xiao, Yining Li

Optical Authenticity in Pushbroom System for Spectral Information Protection.....11320
Pablo Gomez, Roman Jacome, Emmanuel Martinez, Hans Garcia, Henry Arguello

Editing Music with Melody and Text: Using ControlNet for Diffusion Transformer 11325
Siyuan Hou, Shansong Liu, Ruibin Yuan, Wei Xue, Ying Shan, Mangsuo Zhao, Chao Zhang

An Explainable Probabilistic Attribute Embedding Approach for Spoofed Speech Characterization.....11330
Manasi Chhibber, Jagabandhu Mishra, Hye-Jin Shim, Tomi H. Kinnunen

One-Shot Learning for Pose-Guided Person Image Synthesis in the Wild.....11335
Dongqi Fan, Tao Chen, Mingjie Wang, Rui Ma, Qiang Tang, Zili Yi, Qian Wang, Liang Chang

Scaling Multilingual Visual Speech Recognition 11340
K R Prajwal, Sindhu Hegde, Andrew Zisserman

Panorama: An Enabling Technology for Hearables.....11345
Qiyu Rao, Zdenka Babic, Scott C. Douglas, Danilo P. Mandic

LDG: Lightweight Deformable 3D Gaussians for Single-View Dynamic Scene Reconstruction 11350
Youhong Peng, Weixing Xie, Jinwen Li, Shaoqi Wu, Zefeng Wang, Bingbing Hu, Junfeng Yao

Toward Forward-Secure End-To-End Data Sharing: An Attribute-Key-Free CP-ABE Scheme 11355
Xinyi Shi, Yunchuan Guo, Wei Jin, Mingjie Yu, Daiyong Quan, Wenlong Kou, Fenghua Li

Debiased Estimation for Cross-Domain Cold Start Recommendation 11360
Fengxin Li, Hongyan Liu, Jun He, Xiaoyong Du

High-Resolution Speech Restoration with Latent Diffusion Model 11365
Tushar Dhyani, Florian Lux, Michele Mancusi, Giorgio Fabbro, Fritz Hohl, Ngoc Thang Vu

Voice Conversion Via Structural Entropy 11370
Linqin Wang, Zhengtao Yu, Shengxiang Gao, Cunli Mao, Ling Dong, Yuxin Huang

Global Static Pruning Via Adaptive Sample Complexity Awareness 11375
Ming Ma, Yue Wang, Taoli Du, Qinxu Gao, Ying Wang, Wenhui Li

Don't Lose Yourself: Boosting Multimodal Recommendation Via Reducing Node-Neighbor Discrepancy in Graph Convolutional Network 11380
Zheyu Chen, Jinfeng Xu, Haibo Hu

Forensics Analysis of Residual Noise Texture in Digital Images for Detection of Deepfake 11385
Arthur Méreur, Antoine Mallet, Rémi Cogranne, Minoru Kuribayashi

Reinforcement Learning for Charged Particle Beam Control to Minimize Injection Mismatch in Particle Accelerators 11390
Thilina Balasooriya, Shinjae Yoo, Vincent Schoefer, Huan-Hsin Tseng, Yuan Gao, Weijian Lin, Chanaka De Silva

Keeping the Balance: Anomaly Score Calculation for Domain Generalization 11395
Kevin Wilkinghoff, Haici Yang, Janek Ebbers, François G. Germain, Gordon Wichern, Jonathan Le Roux

VoiceGuider: Enhancing Out-Of-Domain Performance in Parameter-Efficient Speaker-Adaptive Text-To-Speech Via Autoguidance 11400
Jiheum Yeom, Heeseung Kim, Jooyoung Choi, Che Hyun Lee, Nohil Park, Sungroh Yoon

Vision Transformers for X-Ray Diffraction Patterns Analysis 11405
Titouan Simonnet, Mame Diarra Fall, Sylvain Grangeon, Bruno Galerne

A Modified Nonlinear Matched Filter for Skewed Noise Based on the Gram-Charlier Expansion 11410
Arie Yeredor

INN-PAR: Invertible Neural Network for PPG to ABP Reconstruction 11414
Soumitra Kundu, Gargi Panda, Saumik Bhattacharya, Aurobinda Routray, Rajlakshmi Guha

Few-Shot Object Detection in Satellite Imagery with Feature Fusion Pyramid and Adaptive Region Proposal Networks 11419
Tuoyu Feng, Weiping Li, Zhijie Tan, Liwen Zhang, Xiang Yuan, Xu Chu

Infrared and Visible Image Fusion with Hierarchical Human Perception 11424
Guang Yang, Jie Li, Xin Liu, Zhusi Zhong, Xinbo Gao

Unsupervised Search for Ethnic Minorities' Medical Segmentation Training Set 11429
Yixiao Chen, Yue Yao, Ruining Yang, Md Zakir Hossain, Ashu Gupta, Tom Gedeon

Enhancing Vision-Language Tracking by Effectively Converting Textual Cues into Visual Cues 11434
Xiaokun Feng, Dailing Zhang, Shiyu Hu, Xuchen Li, Meiqi Wu, Jing Zhang, Xiaotang Chen, Kaiqi Huang

Parametric Object Coding in IVAS: Efficient Coding of Multiple Audio Objects at Low Bit Rates 11439
Andrea Eichenseer, Srikanth Korse, Guillaume Fuchs, Markus Multrus

Self-Supervised Localized Topology Consistency for Noise-Robust Hyperspectral Image Classification 11444
Jie Wang, Liaoyuan Tang, Guanxiong He, Zhe Cao, Zheng Wang, Rong Wang

A Systematic Evaluation of Machine Learning Methods for Fault Detection and Line Identification in Electrical Power Grids.....11449
Julian Oelhaf, Georg Kordowich, Paula Andrea Pérez-Toro, Tomás Arias-Vergara, Andreas Maier, Johann Jäger, Siming Bayer

Multilingual Speaker-Invariant Dysarthria Severity Assessment Using Adversarial Domain Adaptation and Self-Supervised Learning.....11454
Lauren Stumpf, Balasundaram Kadirvelu, A.Aldo Faisal

ConSinger: Efficient High-Fidelity Singing Voice Generation with Minimal Steps 11459
Yulin Song, Guorui Sang, Jing Yu, Chuangbai Xiao

Instantaneous Trajectory Prediction Via Latent Bidirectional Cooperative Diffusion.....11464
Kun Ma, Qilong Han, Jingzheng Yao, Changmao Wu, Chunrui Na

Efficient Global Attention and Correlation-Aware Fusion for Hyperspectral Image Classification.....11469
Hongkang Zhang, Shao-Lun Huang, Ercan Engin Kuruoglu

CAAL-Unet: A Confusion Area Attention Lightweight Network for Brain Tumor Segmentation.....11474
Yang Pu, Qingfeng Wu

Beyond Uniformity: Deblurring Images with Complex Noise Patterns Using Half Quadratic Splitting 11479
Avinash Kumar, Koyyada Dinesh Kumar, Sujit Kumar Sahoo

Beyond Point Annotation: A Weakly Supervised Network Guided by Multi-Level Labels Generated from Four-Point Annotation for Thyroid Nodule Segmentation in Ultrasound Image 11483
Jianning Chi, Zelan Li, Huixuan Wu, Wenjun Zhang, Ying Huang

HCLTS: Mining Customers' Consumption Patterns in Natural Gas Time Series with Hierarchical Contrastive Learning 11488
Yuhang Niu, Jiaqi Ye, Shubao Zhao, Zhaoxiang Hou, Chengyi Yang, Zengxiang Li, Yanlong Wen, Xiaojie Yuan

CASleepNet: A Cross Attention-Based Multimodal Fusion Approach for Sleep Staging with EEG and EOG.....11493
Wei Yu, Jinlong Yang

LaTeXNet: A Specialized Model for Converting Visual Tables and Equations to LaTeX Code 11498
Renqiu Xia, Hongbin Zhou, Ziming Feng, Huanxi Liu, Boan Chen, Bo Zhang, Junchi Yan

Facilitating Semi-Supervised Pedestrian Detection with Structurally Controllable Instance Synthesis.....11503
Tianyou Zhang, Wenhao Wu, Si Wu, Rui Li

CoGAP: A Personalized Federated Learning Method Using Collaborative Optimization for Medical Image Classification 11508
Shenhai Zheng, Congyu Li, Sian Wen, Xi Gao, Lei Yu

Interpolation Filter Design for Sample Rate Independent Audio Effect RNNs 11513
Alistair Carson, Alec Wright, Stefan Bilbao

LagTS: Toward Adaptive Lag Relationship Modeling for Multivariate Time Series Forecasting 11518
Ciyi Liu, Jiaqi Ye, Zhenpeng Yu, Shubao Zhao, Zhaoxiang Hou, Chengyi Yang, Yanlong Wen, Xiaojie Yuan

PDTrack: Progressive Distance Association for Multiple Object Tracking 11523
Yeqiang Liu, Weiran Li, Yanhao Ding, Zhenbo Li

Do Less and Achieve More: Free Condition Video Outpainting with Diffusion Model 11528
Haofan Huang, Yinlin Guo, Yening Lv, Sizhe Shan, Yan Zhang, Yuehai Wang

E-RNS : Enhancing Negative Sample Quality from Gradient Perspective for Graph Recommendation 11533
Qiangsheng Feng, Jiwei Qin, Jie Ma

FedTG: Text-Guided Federated Domain Generalization 11538
Yiming Chen, Nan He, Lifeng Sun

Cross-Lingual Speech Emotion Recognition: Humans Vs. Self-Supervised Models 11543
Zhichen Han, Tianqi Geng, Hui Feng, Jiahong Yuan, Korin Richmond, Yuanchao Li

Local Feature Alignment Prompt-Tuning for Few-Shot Multimodal Aspect Sentiment Analysis 11548
Meirong Ding, Chuang Zou, Hongyi Lin

Distributed Navigation with Dynamic Obstacles 11553
Ellen H. J. Riemens, Raj Thilak Rajan

Continual Unsupervised Domain Adaptation for Audio Deepfake Detection 11558
Xiaohuan Chen, Wenhuan Lu, Ruiteng Zhang, Junhai Xu, Xugang Lu, Lin Zhang, Jianguo Wei

Adversarial Learning for End-To-End Cochlear Speech Denoising Using Lightweight Deep Learning Models 11563
Tom Gajecki, Waldo Nogueira

Effective Context Modeling Framework for Emotion Recognition in Conversations 11568
Cuong Tran Van, Thanh V. T. Tran, Van Nguyen, Truong Son Hy

Design of Multiple Binary Waveforms for the Joint MIMO Radar and Communications 11573
Xiaohan Zhao, Yongzhe Li, Ran Tao

DRANet: Dual-Threshold Guided Reliability Aware Network for Semi-Supervised Image Semantic Segmentation 11578
Lizhe Qiu, Mingyang Zhang, Yuan Zheng

Uncertainty-Aware Crime Prediction with Spatial Temporal Multivariate Graph Neural Networks 11583
Zepu Wang, Xiaobo Ma, Huajie Yang, Weimin Lyu, Yang Liu, Peng Sun, Sharath Chandra Guntuku

OARecon: Object-Aware Viewpoint Augmentation for Indoor Compositional Reconstruction 11588
Yuanyuan Ding, Yiming Fei, Jiandang Yang, Xiaobin Wei, Jiajun Lv, Yong Liu

Debiased Prototype Evolving for Point Cloud Domain Adaptation Via 3D Foundation Models 11593
Feng Yang, Yichao Cao, Xuanpeng Li

Training an Anti-KD Model that Cannot Teach Students Via Similarity Disruption 11598
Zheming Liang, Qi Chu, Tao Gong, Bin Liu, Quanchen Zou, Deyue Zhang, Nenghai Yu

A Hybrid Model for Weakly-Supervised Speech Dereverberation 11603
Louis Bahrman, Mathieu Fontaine, Gaël Richard

Learning a Sparse Polynomial Approximation to the Transition Function of General State-Space Models 11608
Benjamin Cox, Émilie Chouzenoux, Víctor Elvira

Exploring Acoustic Similarity in Emotional Speech and Music Via Self-Supervised Representations 11613
Yujia Sun, Zeyu Zhao, Korin Richmond, Yuanchao Li

Surface Defect Detection Algorithm for Strip Alloy Material Based on Improved YOLOv8 11618
Wei Yang, Jun Yang, Yajin Xia

Camouflaged Object Detection Via Neural Architecture Search 11623
Xin Li, Keren Fu, Qijun Zhao

Sparse-To-Dense Body Surface Potential Mapping Using a Structural Similarity-Enhanced Attention GAN 11628
Ayan Mukherjee, Sawon Pratiher, Oishee Mazumder, Aniruddha Sinha

Energy Consumption Trends in Sound Event Detection Systems 11633
Constance Douwes, Romain Serizel

An Efficient Sample Utilization Method for Deep Learning Based on Class Uncertainty 11638
Jinxin Huang, Qianhua He, Jiezhi Xu, Sam Kwong, Mingru Yang

Conditional Deep Canonical Time Warping 11643
Ran Eisenberg, Afek Steinberg, Ofir Lindenbaum

Leveraging IPA and Articulatory Features as Effective Inductive Biases for Multilingual ASR Training 11648
Jaeyoung Lee, Masato Mimura, Tatsuya Kawahara

Exploiting Beam-Split in IRS-Aided Systems Via OFDMA 11653
P. Siddhartha, L. Yashvanth, Chandra R. Murthy

CAW-CL: Cascaded Adaptive Weighted Contrastive Learning for Unsupervised Ultrasound Plane-Wave Image Reconstruction 11658
Aohua Wang, Jingfeng Lu, Yi Zhang

Calibration of Multiple Asynchronous Microphone Arrays Using Hybrid TDOA 11663
Chengjie Zhang, Wenda Pan, Xinyang Han, He Kong

Multi-View Radar Detection Transformer with Differentiable Positional Encoding 11668
Ryoma Yataka, Pu Perry Wang, Petros Boufounos, Ryuhei Takahashi

Ambisonics Coding in IVAS: A Hybrid SPAR and DirAC System 11673
Dominik Weckbecker, Stefanie Brown, Juan Torres, Markus Multrus, Archit Tamarapu, Guillaume Fuchs

Causal Speech Enhancement with Predicting Semantics Based on Quantized Self-Supervised Learning Features 11678
Emiru Tsunoo, Yuki Saito, Wataru Nakata, Hiroshi Saruwatari

Multi-Objective Large Language Model Unlearning 11683
Zibin Pan, Shuwen Zhang, Yuesheng Zheng, Chi Li, Yuheng Cheng, Junhua Zhao

A Pre-Training Framework that Encodes Noise Information for Speech Quality Assessment 11688
Subrina Sultana, Donald S. Williamson

Lightweight Neural Front-Ends for Low-Resource On-Device Text-To-Speech 11693
Giulia Comini, Heereen Shim, Manuel Sam Ribeiro

Approximation and Analysis of the One-Bit Hermite Law 11698
Chun-Lin Liu, Yi-Hung Chou

Multi-Hop Attention Diffusion Graph Neural Networks for Multimodal Fake News Detection 11703
Zhongqiang Huang, Dongli Lu, Ying Sha

HGNet: Hash Generation Network Guided by High Frequency Information for Fine-Grained Image Retrieval 11708
Hanyun Zhang, Yihua Chen, Xiaoping Liang, Lv Chen, Zhenjun Tang

Weakly Supervised Phonological Features for Pathological Speech Analysis 11713
Jenthe Thienpondt, Geoffroy Vanderreydt, Abdessalem Hammami, Kris Demuynck

A Robust Quality Evaluator for Panoramic Videos 11718
Yi Wang, Jiabao Feng, Yu Zhou, Lijuan Tang, Ruirui Chen, Yanjing Sun, Jicun Ding

SSCM: Self-Supervised Critical Model for Reducing Hallucinations in Chinese Financial Text Generation 11723
Keyan Jin, Yapeng Wang, Leonel Santos, Tao Fang, Xu Yang, Sio Kei Im

Revise, Reason, and Recognize: LLM-Based Emotion Recognition Via Emotion-Specific Prompts and ASR Error Correction 11728
Yuanchao Li, Yuan Gong, Chao-Han Huck Yang, Peter Bell, Catherine Lai

Learning Statistical and Physical Modeling for Consistency Human Motion Prediction 11733
Xunliang Huang, Wenlong Wang, Haoyong Li, Fuming Wang, Cheng Peng, Haoyuan Sun

Generating Customized 4D Motions from Text Inputs Using Spatial-Temporal Slicing Approaches 11738
Zhichao Zhang, Hui Chen, Ming Xu, Jinsheng Deng, Xingshen Song

Exploring Simple Siamese Network for High-Resolution Video Quality Assessment 11743
Guotao Shen, Ziheng Yan, Xin Jin, Longhai Wu, Jie Chen, Ilhyun Cho, Cheul-Hee Hahm

Simultaneous Diarization and Separation of Meetings Through the Integration of Statistical Mixture Models 11748
Tobias Cord-Landwehr, Christoph Boeddeker, Reinhold Haeb-Umbach

MambaFoley: Foley Sound Generation Using Selective State-Space Models 11753
Marco Furio Colombo, Francesca Ronchini, Luca Comanducci, Fabio Antonacci

Optimal One-Hot Logistic Regression for Tree-Based Distribution Classification 11758
Baptiste Schall, Rodolphe Anty, Lionel Fillatre

Complete Reconstruction of the Tongue Contour Through Acoustic to Articulatory Inversion Using Real-Time MRI Data 11763
Sofiane Azzouz, Pierre-André Vuissoz, Yves Laprie

Compositional Audio Representation Learning 11768
Sripathi Sridhar, Mark Cartwright

VP-YOLO: Robust Vehicle-Pedestrian Detection in Challenging Traffic Scenarios Via a Human Visual Perception-Inspired Network 11773
Wenbo Liu, Tao Deng, Fei Yan

Efficient Long-Form Speech Recognition for General Speech In-Context Learning 11778
Hao Yen, Shaoshi Ling, Guoli Ye

Globally Normalizing the Transducer for Streaming Speech Recognition 11783
Rogier C. van Dalen

Polygon Pixel IoU: Similarity Metric Between Polygons with Different Number of Vertices for Arbitrary-Shaped Text Spotting 11788
Rei Endo, Taro Miyazaki, Takahiro Mochizuki, Yoshihiko Kawai

STFEnc: A Novel Deep Encoder for Brain-Computer Interface Based on Interpretable Brain Features 11793
Zonghan Du, Zhongyuan Lai

Acoustic Identification of Individual Animals with Hierarchical Contrastive Learning 11798
Ines Nolasco, Ilyass Moummad, Dan Stowell, Emmanouil Benetos

MPNAS: Multimodal Sentiment Analysis Pruning Via Neural Architecture Search 11803
Binyan Zhang, Ao Ren, Zihao Zhang, Moming Duan, Duo Liu, Yujuan Tan, Kan Zhong

Efficient Fine-Tuning Strategies for Enhancing Face Recognition Performance in Challenging Scenarios 11808
Yin Lin, Ziyang Wu, Qidong Huang, Xinran Liu, Baocai Yin, Jinshui Hu, Bing Yin, Zengfu Wang

ETDE-Net: An End-To-End Time-Domain Enhancement Network for LPI Radar Signals 11813
Chen Cheng, Zhi Sun, Haonan Zhang, Zihao Xiao, Guolong Cui

An Information-Theoretic Analysis of Thompson Sampling with Infinite Action Spaces 11818
Amaury Gouverneur, Borja Rodríguez-Gálvez, Tobias J. Oechtering, Mikael Skoglund

Leveraging Heterophily in Spatial-Temporal Graphs for Multivariate Time-Series Forecasting 11823
Yuxin Chen, Fangru Lin, Jingyi Huo, Hui Yan

Robust Over-The-Air Federated Learning in Heterogeneous Networks 11828
Zubair Shaban, Nazreen Shah, Ranjitha Prasad

Simplified Augmented Real-Valued Time-Delay Neural Network for Digital Predistortion 11833
Lesthuruge Silva, Sri Satish Krishna Chaitanya Bulusu, Nuutti Tervo, Premanandana Rajatheva

A Training-Free Correlation-Weighted Model for Zero-/Few-Shot Industrial Anomaly Detection with Retrieval Augmentation 11838
Wei Ran, Zefang Yu, Suncheng Xiang, Ting Liu, Yuzhuo Fu

Multi-View Image Enhancement Inconsistency Decoupling Guided 3D Gaussian Splatting 11843
Lu Xiao, Jiahao Wu, Zhanke Wang, Guanhua Wu, Runling Liu, Zhiyan Wang, Ronggang Wang

Dual-Population Watermark Vaccine: Efficient and Imperceptible Adversarial Attack for Watermarked Image Protection 11848
Zhili Zhou, Chunhui Zeng, Linna Zhou, Zhongliang Yang, Yujiang Li, Fei Peng, Yong Xie

Semi-Supervised Cognitive State Classification from Speech with Multi-View Pseudo-Labeling 11853
Yuanchao Li, Zixing Zhang, Jing Han, Peter Bell, Catherine Lai

An Experimental Study on Joint Modeling for Sound Event Localization and Detection with Source Distance Estimation 11858
Yuxuan Dong, Qing Wang, Hengyi Hong, Ya Jiang, Shi Cheng

HID-NAS: A Novel Neural Architecture Search Pipeline for High Information Density Data 11863
Wenbo Liu, Tao Deng, Fei Yan

Unified Arbitrary-Time Video Frame Interpolation and Prediction 11868
Xin Jin, Longhai Wu, Jie Chen, Ilhyun Cho, Cheul-Hee Hahm

Minimalistic Video Saliency Prediction Via Efficient Decoder & Spatio Temporal Action Cues 11873
Rohit Girmaji, Siddharth Jain, Bhav Beri, Sarthak Bansal, Vineet Gandhi

Robust Multicomponent Tracking of Ultrasonic Vocalizations 11878
Reyhaneh Abbasi, Nicki Holighaus, Peter Balazs, Vincent Lostanlen, Dustin J. Penn, Sarah M. Zala

Mitigating Motion Artifacts in Non-Contact Respiratory Rate Measurement Utilizing mm-Wave FMCW Radar 11883
Ashi Singhal, Mohd Wajid, Omar Farooq

CAT-Net: A Co-Adaptive Transfer Learning Network for BCI-Assisted Neurorehabilitation 11888
Shuailei Zhang, Yi Ding, Muyun Jiang, Ning Tang, Effie Chew, Kai Keng Ang, Cuntai Guan

ChatCAD: An MLLM-Guided Framework for Zero-Shot CAD Drawing Restoration 11893
Jing Tang, Hongru Xiao, Xiang Li, Wei Wang, Zeyu Gong

MVANet: Multi-Stage Video Attention Network for Sound Event Localization and Detection with Source Distance Estimation 11898
Hengyi Hong, Qing Wang, Ruoyu Wei, Mingqi Cai, Xin Fang

Modulating State Space Model with SlowFast Framework for Compute-Efficient Ultra Low-Latency Speech Enhancement 11903
Longbiao Cheng, Ashutosh Pandey, Buye Xu, Tobi Delbruck, Vamsi Krishna Ithapu, Shih-Chii Liu

Contextual Value Alignment 11908
Pierre Dognin, Jesus Rios, Ronny Luss, Prasanna Sattigeri, Miao Liu, Inkit Padhi, Matthew Riemer, Manish Nagireddy, Kush Varshney, Djallel Bouneffouf

Low-Correlation OFDM Waveform Design with Optimally Coded Sub-Carriers for the Joint Sensing and Communications 11913
Chunxuan Shi, Yongzhe Li, Ran Tao

Generating Targeted Universal Adversarial Perturbation Against Automatic Speech Recognition Via Phoneme Tailoring 11918
Yujun Zhang, Yanqu Chen, Jiakai Wang, Jin Hu, Renshuai Tao, Xianglong Liu

A Weighted Cross-Entropy Loss for Mitigating LLM Hallucinations in Cross-Lingual Continual Pretraining 11923
Yuantao Fan, Ruifan Li, Guangwei Zhang, Chuan Shi, Xiaojie Wang

Multi-Relational Variational Contrastive Learning for Next POI Recommendation 11928
Lin Chen, Peipei Wang, Xiaohui Han, Lijuan Xu

The Missing Piece in Model Editing: A Deep Dive into the Hidden Damage Brought by Model Editing 11933
Jianchen Wang, Zhouhong Gu, Xiaoxuan Zhu, Lin Zhang, Haoning Ye, Zhuozhi Xiong, Sihang Jiang, Hongwei Feng, Yanghua Xiao

Multi-Step Fusion of Relation Type Information and Multi-Task Decoding for Entity Relation Extraction in Ancient Chinese 11938
Chuanlong Tu, Jiali Zuo, Yiyu Hu, Jianyi Wan, Mingwen Wang

Generating Gezi Opera Scores with a Large Language Model and a High-Quality Dataset 11943
Zhen Lei, Ke Gu, Peng Bai, Xiaodong Shi

Efficient Quantization and Denoising Using Local Graph Fourier Frames 11948
Philipp Reingruber, Gerald Matz

Zero-Shot Musical Stem Retrieval with Joint-Embedding Predictive Architectures 11953
Alain Riou, Antonin Gagneré, Gaëtan Hadjeres, Stefan Lattner, Geoffroy Peeters

Quality and Complexity Tradeoffs for DNN-Based Binaural Speech Enhancement in Hearing Aids 11958
Parth Mishra, Deepak Kadetotad, Eric Durant, Terence Betlehem, Martin McKinney

Subpart Suppression Network for Few-Shot Object Counting 11963
Lanxin Liu, Xinyan Liu, Guorong Li

Learning Rank Constrained Exposure Correction from Unpaired Data 11968
Zhuoyue Gong, Qing Zhang

Domain Obfuscation for Efficient Secure Aggregation of Sparse Feature Vectors 11973
Amit Tulsidas Chaulwar

Neural Variational Mode Decomposition and Its Application for ECG Denoising 11978
De-Yan Lu, Jian-Jiun Ding, Yu Tsao

Towards Unbiased Evaluation of Time-Series Anomaly Detector 11983
Debarpan Bhattacharya, Sumanta Mukherjee, Chandramouli Kamanchi, Vijay Ekambaram, Arindam Jati, Pankaj Dayama

Multiple Choice Learning for Efficient Speech Separation with Many Speakers 11988
David Perera, François Derrida, Théo Mariotte, Gaël Richard, Slim Essid

Kernel-Based Anomaly Detection Using Generalized Hyperbolic Processes 11993
Pauline Bourigault, Danilo P. Mandic

EPIC: Error Pattern Informed Correction for Classroom ASR with Limited Labeled Data 11998
Linzhao Jia, Han Sun, Yuang Wei, Changyong Qi, Xiaozhe Yang

An Interactive Evaluation Framework for Empathetic Response Generation 12003
Xixi Lei, Changqun Li, Liang He, Xin Lin

Boundary-Driven Table-Filling with Cross-Granularity Contrastive Learning for Aspect Sentiment Triplet Extraction 12008
Qingling Li, Wushao Wen, Jinghui Qin

TAPO: Task-Referenced Adaptation for Prompt Optimization 12013
Wenxin Luo, Weirui Wang, Xiaopeng Li, Weibo Zhou, Pengyue Jia, Xiangyu Zhao

A Bayesian Perspective on Uncertainty Quantification for Estimated Graph Signals 12018
Lennard Rompelberg, Michael T. Schaub

In Search of Optimal Pretraining Strategy for Robust Speaker Recognition 12023
Nikita Khmelev, Stepan Malykh, Alexandr Anikin, Anastasia Korenevskaya, Sergey Novoselov, Vladimir Volokhov, Zorkina Anastasia, Vladislav Marchevskiy, Galina Lavrentyeva

Audio Decoding by Inverse Problem Solving 12028
Pedro J. Villasana T., Lars Villemoes, Janusz Klejsa, Per Hedelin

Avoiding Domain Drift and Constant Predictions with Diffusion Enhanced Vector-Quantized Autoencoders for Temperature Predictions 12033
Nina Lampl, João Machado de Freitas, Alexander Fuchs, Benedikt Brezina, Michael Klitzsch, Franz Pernkopf

Text-Dependent Speaker Verification Challenge 2024: Exploring Shared and User-Defined Passphrases 12038
Hossein Zeinali, Kong Aik Lee, Jahangir Alam, Lukáš Burget

NAT3DSound: 3D Spatial Sound Field Synthesis with Multi-Modal Non-Autoregressive Transformer 12043
Fuming You, Rongjie Huang, Boyang Zhang, Yongqi Wang, Zhiqing Hong, Zhimeng Zhang, Zhou Zhao

A Novel Multimodal Personality Prediction Method Based on Pretrained Models and Graph Relational Transformer Network 12048
Rongquan Wang, Xianyu Xu, Hao Yang, Lin Wei, Huimin Ma

Perceptual Noise-Masking with Music Through Deep Spectral Envelope Shaping 12053
Clémentine Berger, Roland Badeau, Slim Essid

Deep-Relative-Trust-Based Diffusion for Decentralized Deep Learning 12058
Muyun Li, Aaron Fainman, Stefan Vlaski

Diffusion Based Text-To-Music Generation with Global and Local Text Based Conditioning 12063
Jisi Zhang, Pablo Peso Parada, Md Asif Jalal, Karthikeyan Saravanan

Interactive Robot Action Replanning Using Multimodal LLM Trained from Human Demonstration Videos 12068
Chiori Hori, Motonari Kambara, Komei Sugiura, Kei Ota, Sameer Khurana, Siddarth Jain, Radu Corcodel, Devesh Jha, Diego Romeres, Jonathan Le Roux

INN-Based Secure Steganography Using Lost Information as Adversarial Perturbations 12073
Fei Shang, Weixiang Zhao, Xiangui Kang, Z. Jane Wang

Relaxing Distillation Constraints for Improved New Class Learning in Continual Semantic Segmentation 12078
Zheng Ti, Xuze Hao, Renhai Chen

Prototype Alignment with LoRA Fusion for Class-Incremental Learning 12083
Wei Zhang, Yuan Xie, Zhizhong Zhang, Xin Tan

RadDet: A Wideband Dataset for Real-Time Radar Spectrum Detection..... 12088
Zi Huang, Simon Denman, Akila Pemasiri, Terrence Martin, Clinton Fookes

Convergence Analysis of alpha-SVRG Under Strong Convexity 12093
Sean Xiao, Sangwoo Park, Stefan Vlaski

Learning Rich Speech Representations with Acoustic-Semantic Factorization 12098
Minxue Niu, Najmeh Sadoughi, Abhishek Yanamandra, Pichao Wang, Zhu Liu, Vimal Bhat, S. Elizabeth Norred

Audio Features Investigation for Singing Voice Deepfake Detection 12103
Mahyar Gohari, Davide Salvi, Paolo Bestagini, Nicola Adami

Self-Supervised Learning for Detecting AI-Generated Faces as Anomalies 12108
Mian Zou, Baosheng Yu, Yibing Zhan, Kede Ma

Dual-Frequency Spatio-Temporal Phase Unwrapping 12113
Shuo Du, Qin Zou, Chi Chen, Bisheng Yang

Spatial-Frequency Cross-Domain Mutual Guidance for Visible-Infrared Image Fusion 12118
Zhili Lin, Biao Leng

Metric Learning with Progressive Self-Distillation for Audio-Visual Embedding Learning 12123
Donghuo Zeng, Kazushi Ikeda

TrackNetV4: Enhancing Fast Sports Object Tracking with Motion Attention Maps 12128
Arjun Raj, Lei Wang, Tom Gedeon

Bridging Neural and Symbolic Reasoning: A Dual-System Framework for Interpretable Question Answering 12133
Jihao Shi, Xiao Ding, Hengwei Zhao, Ting Liu, Bing Qin

An Exceptional Dataset for Rare Pancreatic Tumor Segmentation 12138
Wenqi Li, Yingli Chen, Keyang Zhou, Xiaoxiao Hu, Zilu Zheng, Yue Yan, Xinpeng Zhang, Wei Tang, Zhenxing Qian

Voc2vec: A Foundation Model for Non-Verbal Vocalization 12143
Alkis Koudounas, Moreno La Quatra, Sabato Marco Siniscalchi, Elena Baralis

CabiNet: A Deep Learning Framework for Multiclass Medical Image Segmentation from Multiple Single Class Datasets..... 12148
Aman Soni, Ishita Maiti, Nirmalya Ghosh

Learning with Coupled Noisy Labels for Visible-Infrared Person Re-Identification Via Graph Consistency 12153
Zhen Wang, WenXin Zhao, YongFeng Dong

Improving Speech Emotion Recognition in Under-Resourced Languages Via Speech-To-Speech Translation with Bootstrapping Data Selection 12158
Hsi-Che Lin, Yi-Cheng Lin, Huang-Cheng Chou, Hung-yi Lee

Incorporating Spatial Cues in Modular Speaker Diarization for Multi-Channel Multi-Party Meetings 12163
Ruoyu Wang, Shutong Niu, Gaobin Yang, Jun Du, Shuangqing Qian, Tian Gao, Jia Pan

MSANet: Mixed Spectral and Attention Network for Robust 3D Human Pose Estimation 12168
Bing Wang, Suping Wu, Xitie Zhang, Liyuan Shi, Sheng Yang, Zhijian Duan, Tuo Xiong

Correlated Multiple IHC Virtual Staining for Breast Histopathological Images 12173
Xianchao Guan, Zheng Zhang, Yifeng Wang, Ye Zhang, Danling Jiang, Yongbing Zhang

CPL: Curriculum Pseudo Labeling for Weakly Supervised Temporal Forgery Localization 12178
Dijia Zhang, Mingqi Fang, Zhiying Lu, Hongtao Xie

Grouping-Based Crowding Differential Evolution Approaches for Multimodal Feature Selection 12183
Junliu Zhu, Zong-Gan Chen, Jian-Yu Li, Yuncheng Jiang, Zhi-Hui Zhan, Jun Zhang

How Much to Dereverberate? Low-Latency Single-Channel Speech Enhancement in Distant Microphone Scenarios 12188
Satvik Venkatesh, Philip Coleman, Arthur Benilov, Simon Brown, Selim Sheta, Frederic Roskam

3D Mesh Saliency Based on Dictionary Learning with Multi-Level Laplacian-Beltrami Operator 12193
Yu Wang, Xingce Wang, Zhongke Wu, Haichuan Zhao

Node Selection for Physical Layer Security Using Graph-Based Vertex-Frequency Representation 12198
Venugopalachary Kotha, Vijay Kumar Chakka, Shaik Basheeruddin Shah, Nazar T. Ali

Constructing Datasets from Public Police Body Camera Footage 12203
Jamie Rosas-Smith, Martijn Bartelds, Ruizhe Huang, Leibny Paola García-Perera, Karen Livescu, Dan Jurafsky, Anjalie Field

Quant-NeRF: Efficient End-To-End Quantization of Neural Radiance Fields with Low-Precision 3D Gaussian Representation 12208
Ahmed Hasssan, Anupreetham Anupreetham, Jian Meng, Jae-sun Seo

Enhancing Stutter Detection Using Long-Term Average Spectrum Values 12213
Vamshiraghusimha Narasinga, Priyanka Kommagouni, Sridhar Vanga, Kowshik Siva Sai Motepalli, Sai Akarsh C, Purva Barche, Anil Vuppala

XLSR-Transducer: Streaming ASR for Self-Supervised Pretrained Models 12218
Shashi Kumar, Srikanth Madikeri, Juan Zuluaga-Gomez, Esaú Villatoro-Tello, Iuliia Thorbecke, Petr Motlicek, Manjunath K E, Aravind Ganapathiraju

Real-Time Adversarial Attack to Deep Learning-Based Wi-Fi Human Activity Recognition 12223
Byungjun Kim, Amogh Panchagatti, Peter Gerstoft

Speech2rtMRI: Speech-Guided Diffusion Model for Real-Time MRI Video of the Vocal Tract During Speech 12228
Hong Nguyen, Sean Foley, Kevin Huang, Xuan Shi, Tiantian Feng, Shrikanth Narayanan

Constraint-Awareness and Graph Reasoning for Temporal Question Answering 12233
Zheng Sun, Kai Zhang, Xiulong Zhang, Jianting Liu

Enhancing Precision in Image-Guided Spine Surgery Through the Prediction of Occluded Fiducials Utilizing ResNet Architecture 12238
Durga R, Darshan B, Pranav Hyagreev S, Abhilash Chakkaravarthy, Aparna Purayath, Vivek Maik, Manojkumar Lakshmanan, Mohanasankar Sivaprakasam

Adaptive Gradient-Based Timesurface for Event-Based Detection 12243
Ziling Wang, Ziming Wang, Shuang Lian, Rui Yan, Huajin Tang

Weakly-Supervised Video Highlight Detection by Characteristic and Commonality Modeling 12248
Chengze Zhao, Zixuan Zhao, Xu Zhao

Lightweight Multi-Frequency Enhancement Network for RGB-D Video Salient Object Detection 12253
Daerji Suolang, Jiahao He, Wangchuk Tsering, Keren Fu, Xiaofeng Li, Qijun Zhao

Speech Data Selection for Efficient ASR Fine-Tuning Using Domain Classifier and Pseudo-Label Filtering 12258
Pradeep Rangappa, Juan Zuluaga-Gomez, Srikanth Madikeri, Andres Carofilis, Jeena Prakash, Sergio Burdisso, Shashi Kumar, Esaú Villatoro-Tello, Iuliia Nigmatulina, Petr Motlicek, Karthik Pandia, Aravind Ganapathiraju

Open-Vocabulary Visual Emotion Adaptation Via Prompt Learning 12263
Zhaopan Xu, Sicheng Zhao, Xiaojiang Peng, Hongxun Yao

Multi-Scale Re-Weighted Attention Feature Fusion for Non-Intrusive Load Monitoring 12268
Lingxi Yang, Meijun Sun, Haowei Ran, Yipu Liu, Yan Zhou, Zheng Wang

Speech Recognition for Automatically Assessing Afrikaans and isiXhosa Preschool Oral Narratives 12273
Christiaan Jacobs, Annelien Smith, Daleen Klop, Ondřej Klejch, Febe de Wet, Herman Kamper

Precisely Controllable Neural Speech Synthesis 12278
Paul Konstantin Krug, Christoph Wagner, Peter Birkholz, Timo Stich

Dual-Level AMR Injection for Prompt-Based Event Argument Extraction 12283
Xiaojia Huang, Ruifang He, Fei Huang

Age of Gossip with the Push-Pull Protocol 12288
Arunabh Srivastava, Thomas Jacob Maranzatto, Sennur Ulukus

TokenSynth: A Token-Based Neural Synthesizer for Instrument Cloning and Text-To-Instrument 12293
Kyungsu Kim, Junghyun Koo, Sungho Lee, Haesun Joung, Kyogu Lee

Linking Known and Unknown: Generalized Cross-Instance Feature Helps Category Discovery 12298
Yuanhao Zuo, Yichao Liu, Xiwei Liu, Tingzhang Luo

Adaptive Receptive Field Convolution for Top-View Fisheye Images Segmentation 12303
Wenwei Lin, Gang Chen, Changcai Li

Mean Delay in Poisson Wireless Networks with Optimal Packet Fragmentation 12308
S. Vanka, M. Haenggi

Attention-Driven Causal Discovery: From Transformer Matrices to Granger Causal Graphs for Non-Stationary Time-Series Data 12312
Jiageng Zhu, Kehao Li, Zheda Mai, Hanchen Xie, Wael AbdAlmageed, Zubin Abraham

Exploiting Robust Model Watermarking Against the Model Fine-Tuning Attack Via Flat Minima Aware Optimizers 12317
Dongdong Lin, Yue Li, Bin Li, Jiwu Huang

Sketch2Sound: Controllable Audio Generation Via Time-Varying Signals and Sonic Imitations 12322
Hugo Flores García, Oriol Nieto, Justin Salamon, Bryan Pardo, Prem Seetharaman

Robust Deepfake Detection Via Perturbation Domain Alignment 12327
Lin Lu, Yunhong Wang, Liang Zhang, Yuanfang Guo

DFT-Spread-Based OTFS Waveform Design with Good Peak-To-Average Power Ratio for Joint Sensing and Communications 12332
Zhiying Chen, Yongzhe Li, Ran Tao

MMEditor: Multimodal Prompt-Driven 3D Gaussian Splatting Editing 12337
Lihua Lu, Rengang Li, Yaqian Zhao, Xiaohui Zhang, Hui Wei, Ruyang Li

Deep Metamorphic Registration for Tumor-Affected Medical Image Alignment 12342
Wei Jie Pan, Yi Hong

Unsupervised Word Discovery: Boundary Detection with Clustering Vs. Dynamic Programming 12347
Simon Malan, Benjamin van Niekerk, Herman Kamper

Improved Pitch and Voicing Determination Using the Reflected Root Chirp Group Delay Spectrum 12352
Nishant Singh, Mudit Batra, C.S. Ramalingam

NeuroIncept Decoder for High-Fidelity Speech Reconstruction from Neural Activity 12357
Owais Mujtaba Khanday, José L. Pérez-Córdoba, Mohd Yaqub Mir, Ashfaq Ahmad Najar, Jose A. Gonzalez-Lopez

Target Parameter Estimation Using the Capon Method in a MIMO OFDM DFRC System 12362
Satwika Bhogavalli, Eric Grivel, Vincent Corretja

Fractional-Order Hyperbolic Tangent Based Adaptive Algorithm for Feedback Control in Hearing Aids 12367
Vanitha Devi R, Vasundhara, Asutosh Kar, Mads Græsbøll Christensen

EMMeTT: Efficient Multimodal Machine Translation Training 12372
Piotr Żelasko, Zhehuai Chen, Mengru Wang, Daniel Galvez, Oleksii Hrinchuk, Shuoyang Ding, Ke Hu, Jagadeesh Balam, Vitaly Lavrukhin, Boris Ginsburg

Decoupled Feature Matching for Few-Shot Counting and Localization 12377
Chao Zhai, Fan Zhang, Wenyu Chen, Malu Zhang, Fan Li, Xuanting Xie

RF Distillation Diffusion Model: An Efficient RFF Data Augmentation Method 12382
Kai Lin, Caidan Zhao, Jingqian Chen, Liang Xiao

Zero-Shot Image Restoration Via Few-Step Guidance of Consistency Models 12387
Tomer Garber, Tom Tirer

Unsupervised Multi-View Outlier Detection Via Optimal Graph Filtering 12392
Zhiguo Hu, Ning Wang, Peng Zhou, Liang Du, Yuhua Qian, Cheng Wang, Yanming Zhang

Open Automatic Speech Recognition Models for Classical and Modern Standard Arabic 12397
Lilit Grigoryan, Nikolay Karpov, Enas Albasiri, Vitaly Lavrukhin, Boris Ginsburg

Re-Evaluating Privacy in Centralized and Decentralized Learning: An Information-Theoretical and Empirical Study 12402
Changlong Ji, Richard Heusdens, Stephane Maag, Qiongxiu Li

Online Learning from Strategic Human Feedback in LLM Fine-Tuning 12407
Shugang Hao, Lingjie Duan

Prompting DirectSAM for Semantic Contour Extraction in Remote Sensing Images 12412
Shiyu Miao, Delong Chen, Fan Liu, Chuanyi Zhang, Yanhui Gu, Shengjie Guo, Jun Zhou

High-Fidelity Editable Portrait Synthesis with 3D GAN Inversion 12417
Jindong Xie, Jiachen Liu, Yupei Lin, Jinbao Wang, Xianxu Hou, Linlin Shen

DiffListener: Discrete Diffusion Model for Listener Generation 12422
Siyeol Jung, Taehwan Kim

Robust Speech Recognition with Schrödinger Bridge-Based Speech Enhancement 12427
Rauf Nasretdinov, Roman Korostik, Ante Jukić

On the Design of Weakly-Convex Regularizers for Solving Linear Inverse Problems 12432
Abijith Jagannath Kamath, Abhishek Shreekant Bhandiwad, Chandra Sekhar Seelamantula

BEST-STD: Bidirectional Mamba-Enhanced Speech Tokenization for Spoken Term Detection 12437
Anup Singh, Kris Demuynck, Vipul Arora

Identifying Adversarial Attacks in Crowdsourcing Via Dense Subgraph Detection 12442
Abdullah Karaaslanli, Panagiotis A. Traganitis, Aritra Konar

Transfer Learning for Covert Speech Classification Using EEG Hilbert Envelope and Temporal Fine Structure 12447
Saravanakumar Duraisamy, Mateusz Dubiel, Maurice Rekrut, Luis A. Leiva

E-URES 2.0: Efficient User-Centric Residual-Echo Suppression with a Lightweight Neural Network 12452
Amir Ivry, Israel Cohen

Optimal Device Selection and Resource Allocation in Federated Learning 12457
Deepali Kushwaha, Rajesh M. Hegde

Identity-Aware Feature Decoupling Learning for Clothing-Change Person Re-Identification 12462
Haoxuan Xu, Bo Li, Guanglin Niu

Reinforced Domain Selection for Continuous Domain Adaptation 12467
Hanbing Liu, Huaze Tang, Yanru Wu, Yang Li, Xiao-Ping Zhang

Full-Rank No More: Low-Rank Weight Training for Modern Speech Recognition Models 12472
Adriana Fernandez-Lopez, Shiwei Liu, Lu Yin, Stavros Petridis, Maja Pantic

TimeRAG: Boosting LLM Time Series Forecasting Via Retrieval-Augmented Generation 12477
Silin Yang, Dong Wang, Haoqi Zheng, Ruochun Jin

Decentralized Stochastic Successive Convex Approximation for Composite Non-Convex Problems with Non-Linear Functional Constraints 12482
Basil M. Idrees, Shivangi Dubey Sharma, Ketan Rajawat

Subsampling Decomposition Based k-Space Refinement for Accelerated MRI Reconstruction 12487
Xiaoyu Qiao, Weisheng Li, Bin Xiao, Yuping Huang, Lijian Yang

Subset Adaptive Importance Sampling for Multi-Failure-Region Estimation 12492
Sara Helal, Victor Elvira

UMSSS: A Visual Scene Semantic Segmentation Dataset for Underground Mines 12497
Jiawen Wang, Chenfei Liao, Zhongqi Zhao, Lianghui Li, Xuan Gao, Suna Pan, Fangzhen Shi, Yudong Wang, Weijie Zhou, Kehu Yang

LHGNN: Local-Higher Order Graph Neural Networks for Audio Classification and Tagging 12502
Shubhr Singh, Emmanouil Benetos, Huy Phan, Dan Stowell

Tracking Network Dynamics Using Probabilistic State-Space Models 12507
Victor M. Tenorio, Elvin Isufi, Geert Leus, Antonio G. Marques

OSLO-IC: On-The-Sphere Learned Omnidirectional Image Compression with Attention Modules and Spatial Context 12512
Paul Wawerek-López, Navid Mahmoudian Bidgoli, Pascal Frossard, André Kaup, Thomas Maugey

Knowledge Distillation from Ensemble for Spoken Language Identification 12517
Raghuveer Peri, Seyed Omid Sadjadi, Daniel Garcia-Romero, Srikanth Vishnubhotla, Kyu J. Han

Can We "Cherry-Pick"? Investigating Multiple Renditions from a Generative Speech Synthesis Model 12522
Adaeze Adigwe, Sarenne Wallbridge, Zehai Tu, Catherine Lai, Simon King

Distilling Knowledge from Large Video Models for Driver Visual Attention Prediction 12527
Morteza Moradi, Mohammad Moradi, Concetto Spampinato, Ali Borji, Simone Palazzo

Neural Architecture Search for Ultra-Low Memory Blood Glucose Forecasting on the Edge 12532
Hadi Al Zein, Nick van de Waterlaat, Tunc Alkanat

MambaTrack: Exploiting Dual-Enhancement for Night UAV Tracking 12537
Chunhui Zhang, Li Liu, Hao Wen, Xi Zhou, Yanfeng Wang

Multiresolution Encoder-Decoder Convolutional Neural Network for Magnetic Resonance Image Segmentation 12542
Kishore K. Tarafdar, Aaditya Meher, Mirat Shah, Qutubuddin Saifee, Dushyant Kumar, Anant V. Nimkar, Rama Jayasundar, Vikram M. Gadre

REAct: Rational Exponential Activation for Better Learning and Generalization in PINNs 12547
Sourav Mishra, Shreya Hallikeri, Suresh Sundaram

Learning Preconditioners in Gates-Controlled Deep Unfolding Networks Based on Quasi-Newton Methods for Accelerated MRI Reconstruction 12552
Xiaoyu Qiao, Weisheng Li, Bin Xiao, Yuping Huang, Lijian Yang

Preconditioned Sharpness-Aware Minimization: Unifying Analysis and a Novel Learning Algorithm 12557
Yilang Zhang, Bingcong Li, Georgios B. Giannakis

Anchored Monotonic Alignment and Representation Substitution for Rare Spontaneous Behaviors in Spontaneous Speech Synthesis 12562
Ning-Qian Wu, Ya-Jun Hu, Liping Chen, Zhen-Hua Ling

Enabling DMG Wi-Fi Sensing in Data Transmission Intervals by Exploiting Beam Training Codebook 12567
Kareem M. Attiah, Pu Perry Wang, Hassan Mansour, Toshiaki Koike-Akino, Petros Boufounos

Addressing Emotion Ambiguity and Annotator Subjectivity for Enhanced Speech Emotion Labeling 12572
Pooja Kumawat, Aurobinda Routray

Using Ear-EEG to Decode Auditory Attention in Multiple-Speaker Environment 12577
Haolin Zhu, Yujie Yan, Xiran Xu, Zhongshu Ge, Pei Tian, Xihong Wu, Jing Chen

Preference Alignment Improves Language Model-Based TTS 12582
Jinchuan Tian, Chunlei Zhang, Jiatong Shi, Hao Zhang, Jianwei Yu, Shinji Watanabe, Dong Yu

ViCo: A Multitask Video-Enhanced and Cognition-Preserving Modality Alignment Training Framework..... 12587
Zhenda Yu, Jin Chen, Jiayu Shen, Lanxiang Zhou, Han Fang, Xianghao Zang, Chao Ban, Jingfeng Chen, Zhongjiang He, Hao Sun, Zerui Li, Yanmei Kang

Leveraging Multimodal Diffusion Models to Accelerate Imaging with Side Information..... 12592
Timofey Efimov, Harry Dong, Megna Shah, Jeff Simmons, Sean Donegan, Yuejie Chi

IEEE 802.11ad-Aided 5-D Sensing with a UAV Swarm in Urban Environment 12597
Akanksha Sneh, Shobha Sundar Ram, Kumar Vijay Mishra

Towards Robust Category-Level Articulation Pose Estimation Via Integrated Differentiable Rendering 12602
Xinyi Yu, Haonan Jiang, Li Zhang, Yukang Huo, Lin Yuanbo Wu, Yanyan Wei, Rohit Agarwal, Harshal Suresh Shende, Liu Liu, Linlin Ou

Collaborative Semantics-Assisted Large Language Models for Next POI Recommendation..... 12607
Tianci Wang, Yiyuan Wang, Ji Xang

Zero-Resource Speech Translation and Recognition with LLMs..... 12612
Karel Mundnich, Xing Niu, Prashant Mathur, Srikanth Ronanki, Brady Houston, Veera Raghavendra Elluru, Nilaksh Das, Zejiang Hou, Goeric Huybrechts, Anshu Bhatia, Daniel Garcia-Romero, Kyu J. Han, Katrin Kirchhoff

Deep Variational Sequential Monte Carlo for High-Dimensional Observations 12617
Wessel L. van Nierop, Nir Shlezinger, Ruud J.G. van Sloun

Mask-Weighted Spatial Likelihood Coding for Speaker-Independent Joint Localization and Mask Estimation..... 12622
Jakob Kienegger, Alina Mannanova, Timo Gerkmann

AAD-DCE: An Aggregated Multimodal Attention Mechanism for Early and Late Dynamic Contrast Enhanced Prostate MRI Synthesis 12627
Divya Bharti, Sriprabha Ramanarayanan, Sadhana S, Kishore Kumar M, Keerthi Ram, Harsh Agarwal, Ramesh Venkatesan, Mohanasankar Sivaprakasam

Federated Domain Generalization with Label Smoothing and Balanced Decentralized Training..... 12632
Milad Soltany, Farhad Pourpanah, Mahdiyar Molahasani, Michael Greenspan, Ali Etemad

LSTM-QGAN: Scalable NISQ Generative Adversarial Network 12637
Cheng Chu, Aishwarya Hastak, Fan Chen

Injecting Visual Features into Whisper for Parameter-Efficient Noise-Robust Audio-Visual Speech Recognition 12642
Zhao Yang, Yue Heng Yeo, Rui Jiang, Xiao Fu, Weiguang Chen, Wei Xi, Jizhong Zhao

EPE-P: Evidence-Based Parameter-Efficient Prompting for Multimodal Learning with Missing Modalities..... 12647
Zhe Chen, Xun Lin, Yawen Cui, Zitong Yu

Signaling Endothelial Reactivity Induced by Acute Hand Grip: Analyzing Vessel Diameter, Flow Velocity, and EMG Responses..... 12652
Nimmi Sudarsan, Smit Shah, Raj Kiran V, P M Nabeel, Dinu S Chandran, Jayaraj Joseph

WiSenseNet: A Unified Foundation Model for Diverse Wi-Fi Sensing Tasks Using Channel State Information 12657
Niall Lyons, Ashutosh Pandey, Avik Santra

Speech Separation Based on Pre-Trained Model and Deep Modularization 12662
Peter Ochieng

Diversity Matters: Co-Training for Semi-Supervised Change Detection in Remote Sensing Images 12667
Zan Mao, Xin Li, Ze Luo, Yingchao Piao

Compgen: Synthesis and Generation of Faces from Edgemaps 12672
Ruban Vishnu Pandian, Abhiram Rao Gorle, Karthick Krishna M, Nambi Seshadri, R. David Koilpillai

Exploring Facial Kinship Verification Through Contactless Heart Activity Analysis 12677
Xiaoting Wu, Xiaoyi Feng, Constantino Álvarez Casado, Lili Liu, Miguel Bordallo López

BANC: Towards Efficient Binaural Audio Neural Codec for Overlapping Speech 12682
Anton Ratnarajah, Shi-Xiong Zhang, Dong Yu

Open-Modality Latent Modality Interaction Maximization for Audio-Visual Learning 12687
Zhao Yang, Rui Jiang, Xiao Fu, Wei Xi, Jizhong Zhao

Object-Centric Discriminative Learning for Text-Based Person Retrieval 12692
Haiwen Li, Delong Liu, Fei Su, Zhicheng Zhao

Trustworthy Recommendation for Consumer Electronics Using Hypernetworks 12697
Yicheng Di, Song Shen, Jiayu Bao, Yuan Liu

Digital Twin-Driven Bearing-Fault Detection in Induction Motor and Drives Using Graph Sampling and Aggregation Network 12702
Haraprasad Badajena, Suryanarayan Majhi, Bivash Chakraborty, Mamata Jenamani, Aurobinda Routray, Ronit Dutta

Effective Techniques for Scaling Audio Encoder Pretraining 12707
Byeonggeun Kim, Andrew Bydlon, Qingming Tang, Huy Phan, Chieh-Chi Kao, Tao Zhang, Chao Wang

Sequence Knowledge Enhancement Distillation Framework for Ultra-Fast Image Deraining 12712
Jihao Li, Jincheng Hu, Ming Liu, Pengyu Fu, Jingjing Jiang, Yuanjian Zhang

Fully Spiking Neural Network for Legged Robots 12717
Xiaoyang Jiang, Qiang Zhang, Jingkai Sun, Jiahang Cao, Jingtong Ma, Renjing Xu

Online Scalable Gaussian Processes with Conformal Prediction for Guaranteed Coverage 12722
Jinwen Xu, Qin Lu, Georgios B. Giannakis

CiGA: A Cross-Layer Fine-Grained Attention Correction Method for Large Language Model 12727
Duo Li, Jing Zhao

A Unified Metric for Simultaneous Evaluation of Error Rate and Annotation Cost 12732
Mark Lindsey, Francis Kubala, Richard M. Stern

Outage Analysis of IRS-Aided Wireless Energy Transfer Under Correlation and Imperfect CSI 12737
Chandan Kumar, Salil Kashyap

Model-Driven Learning Approach for Robust WiFi-Based Fall Detection 12742
Sai Deepika Regani, Beibei Wang, K. J. Ray Liu

SPEA: Large-Scale Entity Alignment Via Self-Partitioning 12747
Weiguo Chen, Changjian Wang, Kele Xu, Yuan Yuan, Wei Chen, Zixuan Dong

Metadata-Assisted Pose Correction for Immersive Audio Split Rendering 12752
Rishabh Tyagi, Stefan Bruhn, Jeroen Breebaart

Dual-Path Mamba: Short and Long-Term Bidirectional Selective Structured State Space Models for Speech Separation 12757
Xilin Jiang, Cong Han, Nima Mesgarani

CPT-Boosted Wav2vec2.0: Towards Noise Robust Speech Recognition for Classroom Environments 12762
Ahmed Adel Attia, Dorottya Demszky, Tolúlọpe Ògúnremí, Jing Liu, Carol Espy-Wilson

GMMCL: Adaptive Concept Drift in Data Streams with Gaussian Mixture Models Based on Contrastive Learning 12767
Hongwei Wu, Jin Pan, Rong Yang, Hong Zhang, Guang Shi, Zhuojun Jiang, Qingyun Liu

GDRIVE: Adaptive Object Detection in Autonomous Vehicles Via Graph-Based Feature Learning 12772
SuYang Xi, Yunhao Liu, Hao Lu, Yi Ding, Hong Ding

Implementing Finite Impulse Response Filters on Quantum Computers 12777
Aishwarya Majumdar, Bojko N. Bakalov, Dror Baron, Yuan Liu

M*: On-Chip Microfluidic Operations with A-Star for Portable Diagnostics 12782
Cheng-Hsuan Hsieh, Eugene Lee, Peng Hsien Chi, Jiajie Diao, Chen-Yi Lee

A Multi-Modal Information Fusion Model for Automatic Sleep Staging 12787
Xuhui Wang, Yuanyuan Zhu, Xiaodong Jia

Investigating the Sensitivity of Pre-Trained Audio Embeddings to Common Effects 12792
Victor Deng, Changhong Wang, Gaël Richard, Brian McFee

TKA-MIL: Top-K Attention Multiple Instance Learning for Whole Slide Image Classification and Instance Probability Derivation 12797
Sicheng Yu, Xingshu Chen, Fangzhou Cao, Ting Tian

Optimization of Beamwidth in Automotive Radars Based on Statistics of Street Geometry 12802
Mohammad Taha Shah, Gourab Ghatak, Shobha Sundar Ram

LLM Supervised Pre-Training for Multimodal Emotion Recognition in Conversations 12807
Soumya Dutta, Sriram Ganapathy

Boolean Matrix Compressed Sensing 12812
Qiang Liu, Mahdi Soleymani, Hessam Mahdavifar

Metadata-Assisted Spatial Audio Coding in IVAS Codec 12817
Adriana Vasilache, Tapani Pihlajakuja, Mikko-Ville Laitinen

Temporally Aligned Audio for Video with Autoregression 12822
Ilpo Viertola, Vladimir Iashin, Esa Rahtu

A Riemannian Approach to Ground Metric Learning for Optimal Transport 12827
Pratik Jawanpuria, Dai Shi, Bamdev Mishra, Junbin Gao

Towards a Single ASR Model that Generalizes to Disordered Speech 12832
Jimmy Tobin, Katrin Tomanek, Subhashini Venugopalan

Transcribing and Translating, Fast and Slow: Joint Speech Translation and Recognition 12837
Niko Moritz, Ruiming Xie, Yashesh Gaur, Ke Li, Simone Merello Zeeshan Ahmed, Frank Seide, Christian Fuegen

Sparse Generation: Making Pseudo Labels Sparse for Point Weakly Supervised Object Detection on Low Data Volume 12842
Chuyang Shang, Tian Ma, Wanzhu Ren, Yuancheng Li, Jiayi Yang

Semantic Graph Embedded Energy Minimization Learning for Scene Graph Generation 12847
Jinghang Chen, Chi Zhang, Yuehu Liu, Le Wang

End-To-End Acoustic-Articulatory Dysarthric Speech Recognition Leveraging Large-Scale Pretrained Acoustic Features 12852
Zhengjun Yue, Yuanyuan Zhang

Audio Diffusion with Large Language Models 12857
Yinghui Huang, Kyle Kastner, Kartik Audhkhasi, Bhuvana Ramabhadran, Andrew Rosenberg

Concentrating Harder for Faster Audio Transformer 12862
Lorenz P. Schmidt, Nils Peters

Higher-Order Topological Directionality and Directed Simplicial Neural Networks 12867
Manuel Lecha, Andrea Cavallo, Francesca Dominici, Elvin Isufi, Claudio Battiloro

HRTF Estimation Using a Score-Based Prior 12872
Etienne Thuillier, Jean-Marie Lemercier, Eloi Moliner, Timo Gerkmann, Vesa Välimäki

Keeping the Best: The K-Best Rule for Efficient Quickest Change Detection with Unknown Post-Change Distribution 12877
James Zachary Hare, Lance Kaplan, Venugopal V. Veeravalli, Don Towsley

Driver Reaction Time Prediction Through Adaptive Evolutionary Synchrony Window and Convolutional-LSTM 12882
Adarsh V Parekkattil, Vivek Singh, Sanjeev Kumar Varun, Tharun Kumar Reddy Bollu

Cross-Lingual Evaluation of Hypernasality Using Wav2Vec2 Features 12887
Krupaben Kothadia, Vikram C M, Ajish Abraham, Pushpavathi M, S R Mahadeva Prasanna, Nancy Scherer, Kathy Chapman, Julie Liss, Visar Berisha

MFMamba: A Multimodal Fusion State Space Model for Depression Recognition 12892
Jingyi Liu, Yuanyuan Shang, Mengyuan Yang, Zhuhong Shao, Jiaxi Lu, Tie Liu

Enhancing Standard and Dialectal Frisian ASR: Multilingual Fine-Tuning and Language Identification for Improved Low-Resource Performance 12897
Reihaneh Amooie, Wietse De Vries, Yun Hao, Jelske Dijkstra, Matt Coler, Martijn Wieling

3GPP IVAS Codec – Perspectives on Development, Testing and Standardization 12902
S. Bruhn, T. Toftgård, S. Döhla, H-Y. Su, L. Laaksonen, T. Moriya, S. Ragot, H. Ehara, M. Szczerba, I. Varga, A. Schevciw, M. Jelinek

Graph Neural Networks Meet Probabilistic Graphical Models: A Survey 12907
Chenqing Hua, Sitao Luan, Qian Zhang, Jie Fu, Guy Wolf

Multivariate Time Series Data Mining for Failure Prediction & Root Cause Analysis 12912
Naman Agarwal, Gagan Raj Gupta, Vishwesh Jatala, Aniket Saha

Modeling of HR Filters for Audio Object Rendering 12917
Sumeyra Demir Kanik, Erlendur Karlsson, Tomas Toftgard, Erik Norvell

Efficient MDCT-Based Multi-Channel Coding with Perceptual Whitening and Broadband ILD Compensation 12922
Goran Marković, Eleni Fotopoulou, Jan Frederik Kiene, Christian R. Helmrich

Optimized Self-Supervised Training with BEST-RQ for Speech Recognition 12927
Ilja Baumann, Dominik Wagner, Korbinian Riedhammer, Tobias Bocklet

PAIR: Complementarity-Guided Disentanglement for Composed Image Retrieval 12932
Zhiheng Fu, Zixu Li, Zhiwei Chen, Chunxiao Wang, Xuemeng Song, Yupeng Hu, Liqiang Nie

A First-Order DirAC-Based Parametric Ambisonic Coder for Immersive Communications 12937
Guillaume Fuchs, Florin Ghido, Dominik Weckbecker, Oliver Thiergart

Speech Recognition with LLMs Adapted to Disordered Speech Using Reinforcement Learning 12942
Chirag Nagpal, Subhashini Venugopalan, Jimmy Tobin, Marilyn Ladewig, Katherine Heller, Katrin Tomanek

Multi-Level Feature Adaptation and Embeddings Alignment for Zero-Shot Anomaly Detection 12947
Yiqing Liu, Huilin Deng, Gang Zhao

Speech Slytherin: Examining the Performance and Efficiency of Mamba for Speech Separation, Recognition, and Synthesis 12952
Xilin Jiang, Yinghao Aaron Li, Adrian Nicolas Florea, Cong Han, Nima Mesgarani

Retrieval Augmented Correction of Named Entity Speech Recognition Errors 12957
Ernest Pusateri, Anmol Walia, Anirudh Kashi, Bortik Bandyopadhyay, Nadia Hyder, Sayantan Mahinder, Raviteja Anantha, Daben Liu, Sashank Gondala

M-BEST-RQ: A Multi-Channel Speech Foundation Model for Smart Glasses 12962
Yufeng Yang, Desh Raj, Ju Lin, Niko Moritz, Junteng Jia, Gil Keren, Egor Lakomkin, Yiteng Huang, Jacob Donley, Jay Mahadeokar, Ozlem Kalinli

Single-Stage TTS with Masked Audio Token Modeling and Semantic Knowledge Distillation 12967
Gerard I. Gállego, Roy Fejgin, Chunghsin Yeh, Xiaoyu Liu, Gautam Bhattacharya

A Non-Autoregressive Model for Joint STT and TTS 12972
Vishal Sunder, Brian Kingsbury, George Saon, Samuel Thomas, Slava Shechtman, Hagai Aronowitz, Eric Fosler-Lussier, Luis Lastras

PromptSeg: Learning to Segment Medical Image Via Visual Prompts 12977
Minfan Zhao, Ziqi Zhu, Jun Shi, Zhaohui Wang, Junshi Chen, Hong An, Bing Yan

MedFocusCLIP: Improving Few Shot Classification in Medical Datasets Using Pixel Wise Attention 12982
Aadya Arora, Vinay Namboodiri

Robust Hybrid Beamforming for Integrated Sensing and Communications Via Learned Optimization 12987
Lei Wang, Sergiy A. Vorobyov, Esa Ollila

Textless Streaming Speech-To-Speech Translation Using Semantic Speech Tokens 12992
Jinzheng Zhao, Niko Moritz, Egor Lakomkin, Ruiming Xie, Zhiping Xiu, Katerina Zmolikova, Zeeshan Ahmed, Yashesh Gaur, Duc Le, Christian Fuegen

HR-SKGs: Hyper-Relational Semantic Knowledge Graphs for Multi-Hop Reading Comprehension 12997
Yingying Zhang, Bo Cheng, Yuli Chen

Multi-Agent Reinforcement Learning in Partially Observable Environments Using Social Learning..... 13002
Ainur Zhaikhan, Ali H. Sayed

Strong Alone, Stronger Together: Synergizing Modality-Binding Foundation Models with Optimal Transport for Non-Verbal Emotion Recognition 13006
Orchid Chetia Phukan, Mohd Mujtaba Akhtar, Girish, Swarup Ranjan Behera, Sishir Kalita, Arun Balaji Buduru, Rajesh Sharma, S.R Mahadeva Prasanna

First-Frame Supervised Video Polyp Segmentation Via Propagative and Semantic Dual-Teacher Network 13011
Qiang Hu, Mei Liu, Qiang Li, Zhiwei Wang

Semi-Supervised Contrastive Learning for Controllable Video-To-Music Retrieval 13016
Shanti Stewart, Gouthaman KV, Lie Lu, Andrea Fanelli

Diffusion Model Based Image Reconstruction in Lensless Imaging..... 13021
Ashish Verma, Vivek Boominathan, Ashok Veeraraghavan, Chandra Sekhar Seelamantula

Exploring Prediction Targets in Masked Pre-Training for Speech Foundation Models 13026
Li-Wei Chen, Takuya Higuchi, He Bai, Ahmed Hussen Abdelaziz, Shinji Watanabe, Alexander Rudnicky, Tatiana Likhomanenko, Barry-John Theobald, Zakaria Aldeneh

Domain Connection Based Unsupervised Domain Adaptation for Semantic Segmentation..... 13031
Chunze Yang, Xiaodong Zhang, Peiyuan Tang, Haoran Yuan, Haojie Xin, Zijiang James Yang

Multi-Label Recognition Under Noisy Supervision: A Confusion Mixture Modeling Approach 13036
Diego Linares Gonzalez, Shahana Ibrahim

Transducer-Llama: Integrating LLMs into Streamable Transducer-Based Speech Recognition..... 13041
Keqi Deng, Jinxi Guo, Yingyi Ma, Niko Moritz, Philip C. Woodland, Ozlem Kalinli, Mike Seltzer

Single Trial Reaction Time Prediction Using Optimal Synchrony Window Detection..... 13046
Adarsh V Parekkattil, Sanjeev Kumar Varun, Tharun Kumar Reddy Bollu

Generalizable Real-Time Accelerated Dynamic MRI 13051
Silpa Babu, Wahidul Alam, Rushdi Zahid Rusho, Sajan Goud Lingala, Namrata Vaswani

Hierarchical Expectation Propagation for Semi-Blind Channel Estimation in Cell-Free Networks 13056
Zilu Zhao, Dirk Slock

Codar: Complex-Valued Neural Network for Crossing-Floor Intrusion Detection Via WiFi..... 13061
Weiting Ou, Yipeng Liu, Zhijie Sun, Bing Li, Le Zhang, Ce Zhu

Prediction-Driven Untrained Network for Single-Snapshot Sparse Array Interpolation..... 13066
Yinyan Bu, Jiajie Yu, Piya Pal

Dynamic Incentive Model for Federated Learning Model Trading Via Evolutionary Game Theory..... 13071
Tianxiang Chen, Feng Wang, Wenjie Hou, Shaoting Tang, Zhiming Zheng

SepMamba: State-Space Models for Speaker Separation Using Mamba 13076
Thor Højhus Avenstrup, Boldizsár Elek, István László Mádi, András Bence Schin, Morten Mørup, Bjørn Sand Jensen, Kenny Olsen

Audio-Visual Representation Learning for Lip-Sync Estimation Through Ranking Augmented Contrastive Training 13081
Bhavin Jawade, Ravi Teja Gadde, Christophe Bejjani, Yinghong Lan

Towards Detecting Auditory Attention from in-Ear Muscle Contractions Using Commodity Earbuds 13086
Harshvardhan Takawale, Yang Liu, Khaldoon Al-Naimi, Fahim Kawsar, Alessandro Montanari

StereoMamba: Enhancing Stereo Image Super-Resolution with Structured State Space Models and Bi-Directional Cross Attention 13091
Zhenchao Ma, Hamid Reza Tohidypour, Panos Nasiopoulos, Victor C. M. Leung

Towards Time-Frequency Deformation Stability Bounds for Deep Convolutional Neural Networks 13096
Albert Chua

DynamicAttention: Dynamic KV Cache for Disaggregate LLM Inference 13101
Zhiqiang Ding, Tongkai Yang

Speaker Embedding Informed Audiovisual Active Speaker Detection for Egocentric Recordings 13106
Jason Clarke, Yoshihiko Gotoh, Stefan Goetze

The Evolution of System on Chip Integrated Circuits for Hearing-Aid Signal Processing..... 13111
James Ryan, Ivo Coenen, Robert Brennan

Uniform Convergence of Lipschitz Functions with Dependent Gaussian Samples 13116
Mina Sadat Mahmoudi, Saeed Foroutan, Seyed Abolfazl Motahari, Babak Khalaj

Investigating Voiced and Unvoiced Regions of Speech for Audio Deepfake Detection 13121
Ganesh Sivaraman, Hemlata Tak, Elie Khoury

Self-Supervised Learning-Based Multimodal Prediction on Prosocial Behavior Intentions 13126
Abinay Reddy Naini, Zhaobo K. Zheng, Teruhisa Misu, Kumar Akash

Some Intriguing Observations on the Learnt Matrices in Deep Unfolded Networks 13131
Kartheek Kumar Reddy Nareddy, Inbasekaran Perumal, Chandra Sekhar Seelamantula

Pruning Then Reweighting: Towards Data-Efficient Training of Diffusion Models 13136
Yize Li, Yihua Zhang, Sijia Liu, Xue Lin

MARS6: A Small and Robust Hierarchical-Codec Text-To-Speech Model 13141
Matthew Baas, Pieter Scholtz, Arnav Mehta, Elliott Dyson, Akshat Prakash, Herman Kamper

Model-Based Online Millimeter-Wave Channel Sensing with Learned Empirical Priors..... 13146
Parthasarathi Khirwadkar, Bhaskar D. Rao, Piya Pal

Mamba Meets Financial Markets: A Graph-Mamba Approach for Stock Price Prediction..... 13151
Ali Mehrabian, Ehsan Hoseinzade, Mahdi Mazloum, Xiaohong Chen

Sample Efficient Reinforcement Learning Via Large Vision Language Model Distillation..... 13156
Donghoon Lee, Tung M. Luu, Younghwan Lee, Chang D. Yoo

MEDIAN: Adaptive Intermediate-Grained Aggregation Network for Composed Image Retrieval 13161
Qinlei Huang, Zhiwei Chen, Zixu Li, Chunxiao Wang, Xuemeng Song, Yupeng Hu, Liqiang Nie

Reward Generation Via Large Vision-Language Model in Offline Reinforcement Learning 13166
Younghwan Lee, Tung M. Luu, Donghoon Lee, Chang D. Yoo

Cognitive MIMO Radar Beamforming for Target Tracking Using a BCRB-Based Criterion 13171
Helin Sun, Joseph Tabrikian, Hagit Messer, Hongyuan Gao

Improving Pronunciation and Accent Conversion Through Knowledge Distillation and Synthetic Ground-Truth from Native TTS 13176
Tuan Nam Nguyen, Seymanur Akti, Ngoc Quan Pham, Alexander Waibel

Contrastive Knowledge Distillation for Embedding Refinement in Personalized Speech Enhancement 13181
Thomas Serre, Mathieu Fontaine, Éric Benhaim, Slim Essid

Text-Infused Audio-Visual Video Parsing with Semantic-Aware Multimodal Contrastive Learning 13186
Pengcheng Zhao, Yanxiang Chen, Dan Guo, Yuanzhi Yao

A Zero-Shot Physics-Informed Dictionary Learning Approach for Sound Field Reconstruction 13191
Stefano Damiano, Federico Miotello, Mirco Pezzoli, Alberto Bernardini, Fabio Antonacci, Augusto Sarti, Toon Van Waterschoot

Robust Sensor Selection by Deep Unfolding 13196
Yuvraj Singh, Jahnvi Rohela, Kaushani Majumder, Satish Mulleti

Masked Latent Prediction and Classification for Self-Supervised Audio Representation Learning 13201
Aurian Quelennec, Pierre Chouteau, Geoffroy Peeters, Slim Essid

Wasserstein Heterogeneous Graph Neural Networks for Uncertainty-Aware Anomaly Detection 13206
Chen Chen, Yunchun Li, Boxuan Jiao, Guorui Zhao, Wei Li

On Overlap Ratio in Defocused Electron Ptychography 13211
Amirafshar Moshtaghpour, Angus I. Kirkland

Role of the Pretraining and the Adaptation Data Sizes for Low-Resource Real-Time MRI Video Segmentation 13216
Masoud Thajudeen Tholan, Vinayaka Hegde, Chetan Sharma, Prasanta Kumar Ghosh

Distillation and Pruning for Scalable Self-Supervised Representation-Based Speech Quality Assessment 13221
Benjamin Stahl, Hannes Gamper

XAI for Gender Representation in Media Analysis 13226
François Buet, Camille Guinaudeau, Cyril Grouin, Sahar Ghannay, Shin'Ichi Satoh

Adaptive Acquisition in Bayesian Optimization with Agnostic Ensembles 13231
Anand Ravishankar, Fernando Llorente, Yuanqing Song, Petar M. Djurić

Subject Representation Learning from EEG Using Graph Convolutional Variational Autoencoders 13236
Aditya Mishra, Ahnaf Mozib Samin, Ali Etemad, Javad Hashemi

Mamba-Based Segmentation Model for Speaker Diarization 13241
Alexis Plaquet, Naohiro Tawara, Marc Delcroix, Shota Horiguchi, Atsushi Ando, Shoko Araki

Diffusion Learning Over Adaptive Competing Networks 13246
Yike Zhao, Haoyuan Cai, Ali H. Sayed

TSIformer: Multi-Scale Dilation Transformer with Cross-Variable and Cross-Feature Dependency for Time Series Imputation 13251
Haozheng Yang, Xuelin Cheng, Runjie Zhao, Xu Zou, Botao Wu, Xince Chen

Improved Image Classification with Manifold Neural Networks 13256
Caio F. Deberaldini Netto, Zhiyang Wang, Luana Ruiz

Adaptive Central Frequencies Locally Competitive Algorithm for Speech 13261
Soufiyan Bahadi, Eric Plourde, Jean Rouat

Spy Inside: Scalable Verification of Dependable Transformers for Event Time Series Systems 13266
Haodong Deng, Qi Qi, Lu Lu, Zirui Zhuang, Xingyu Zeng, Jinguang Wang, Bo He, Wei Li, Jingyu Wang

SpeechTaxi: On Multilingual Semantic Speech Classification 13271
Lennart Keller, Goran Glavaš

Robust Frame-Level Speaker Localization in Reverberant and Noisy Environments by Exploiting Phase Difference Losses 13276
Shanmukha Srinivas Battula, Hassan Taherian, Ashutosh Pandey, Daniel Wong, Buye Xu, DeLiang Wang

FDR-Controlled Portfolio Optimization for Sparse Financial Index Tracking 13281
Jasin Machkour, Daniel P. Palomar, Michael Muma

Frequency-Based Federated Domain Generalization for Polyp Segmentation 13286
Hongyi Pan, Debesh Jha, Koushik Biswas, Ulas Bagci

Support Recovery in 1-Bit Compressed Sensing with Burst Sparse Noise 13291
Saikiran Bulusu, Venkata Gandikota, Pramod K. Varshney

Adaptive Compression of Supervised and Self-Supervised Models for Green Speech Recognition 13296
Mouaad Oujabour, Leila Ben Letaifa, Jean-François Dollinger, Jean-Luc Rouas

Revisiting Acoustic Features for Robust ASR 13301
Muhammad A. Shah, Bhiksha Raj

Novel Deep Gaussian Process Structures with Flexible Depths 13306
Yuanqing Song, Yuhao Liu, Petar M. Djurić

Pre-Training, Fine-Tuning and Re-Ranking: A Three-Stage Framework for Legal Question Answering 13311
Shiwen Ni, Hao Cheng, Min Yang

Leveraging Mixture of Experts for Improved Speech Deepfake Detection 13316
Viola Negroni, Davide Salvi, Alessandro Ilic Mezza, Paolo Bestagini, Stefano Tubaro

Acoustic Position Estimation of a Silent Listener 13321
Jeremy Lawrence, Cagdas Tuna, Andreas Walther, Nils Peters

Sequential Contrastive Audio-Visual Learning 13326
Ioannis Tsiamas, Santiago Pascual, Chunghsin Yeh, Joan Serrà

Subjective Quality Evaluation of Point Clouds Using a Head-Mounted Display 13331
Joao Prazeres, Rafael Rodrigues, Manuela Pereira, Antonio M. G. Pinheiro

DiffSSD: A Diffusion-Based Dataset for Speech Forensics 13336
Kratika Bhagtani, Amit Kumar Singh Yadav, Paolo Bestagini, Edward J. Delp

SpectralCam: High-Resolution Low-Cost Spectral Imaging Using DSLR Cameras 13341
A. Paruchuri, A. Ramirez-Jaime, G. Arce, A. Alrushud, Xu Ma, R. Radpour

A Novel Underwater Acoustic Signal Denoising Model Based on Complex Convolution Dual-Branch Multi-Scale Attention Network 13346
Jianxun Tang, Zhe Chen, Mingsong Chen

Analysis of Speech Temporal Dynamics in the Context of Speaker Verification and Voice Anonymization 13351
Natalia Tomashenko, Emmanuel Vincent, Marc Tommasi

Bayesian Filtering on Graphs 13356
Bishwadeep Das, Madeline Navarro, Santiago Segarra, Elvin Isufi

Learning Music Audio Representations with Limited Data 13361
Christos Plachouras, Emmanouil Benetos, Johan Pauwels

Learning Control of Neural Sound Effects Synthesis from Physically Inspired Models..... 13366
Yisu Zong, Joshua Reiss

The Misspecified Cramér-Rao Bound for DOA Estimation with One-Bit Quantized Data 13371
Nadav E. Rosenthal, Joseph Tabrikian

Implicit Neural Representations with Fourier Kolmogorov-Arnold Networks 13376
Ali Mehrabian, Parsa Mojarad Adi, Moein Heidari, Ilker Hacihaliloglu

DreamVideo: High-Fidelity Image-To-Video Generation with Image Retention and Text Guidance..... 13381
Cong Wang, Jiaxi Gu, Panwen Hu, Yuanfan Guo, Xiao Dong, Hang Xu, Xiaodan Liang

EasyControl: Adding Control to Video Diffusion for Controllable Video Generation and Interpolation 13386
Cong Wang, Jiaxi Gu, Panwen Hu, Xiao Dong, Yuanfan Guo, Hang Xu, Xiaodan Liang

Second-Order Wireless Federated Leaning Via Nonparametric Hessian Estimation 13391
Shayan Mohajer Hamidi, Ali Bereyhi

Detecting Neurodegenerative Diseases Using Frame-Level Handwriting Embeddings 13396
Sarah Laouedj, Yuzhe Wang, Jesús Villalba, Thomas Thebaud, Laureano Moro-Velázquez, Najim Dehak

Harnessing the Potential of Omnidirectional UAVs in RIS-Enabled Wireless Networks 13401
Abdoul Karim A. H. Saliah, Hajar El Hammouti, Daniel Bonilla Licea

Effective Integration of KAN for Keyword Spotting 13406
Anfeng Xu, Biqiao Zhang, Shuyu Kong, Yiteng Huang, Zhaojun Yang, Sangeeta Srivastava, Ming Sun

Dynamic Prototype Rehearsal for Continual ECG Arrhythmia Detection 13411
Sana Rahmani, Reetam Chatterjee, Ali Etemad, Javad Hashemi

Towards HRTF Personalization Using Denoising Diffusion Models 13416
Juan Camilo Albarracín Sánchez, Luca Comanducci, Mirco Pezzoli, Fabio Antonacci

SEE: Semantically Aligned EEG-To-Text Translation 13421
Yitian Tao, Yan Liang, Luoyu Wang, Yongqing Li, Qing Yang, Han Zhang

LLM-Augmented Symbolic RL with Landmark-Based Task Decomposition 13426
Alireza Kheirandish, Duo Xu, Faramarz Fekri

Rethinking Decoding in Multi-Intent Spoken Language Understanding 13431
Ying Xia, Zhen Xiong, Kefan Shen, Zhihong Zhu, Shaorong Xie, Wei Liu

A Quantitative Metric Selection Approach for Time-Series Forecasting Foundation Models 13436
Hongjie Chen, Akshay Mehra, Josh Kimball, Sungchul Kim

Get Large Language Models Ready to Speak: A Late-Fusion Approach for Speech Generation 13441
Maohao Shen, Shun Zhang, Jilong Wu, Zhiping Xiu, Ehab AlBadawy, Yiting Lu, Mike Seltzer, Qing He

ImmerseDiffusion: A Generative Spatial Audio Latent Diffusion Model 13446
Mojtaba Heydari, Mehrez Souden, Bruno Conejo, Joshua Atkins

Speech Re-Painting for Robust ASR 13451
Kyle Kastner, Gary Wang, Isaac Elias, Takaaki Saeki, Pedro Moreno Mengibar, Françoise Beaufays, Andrew Rosenberg, Bhuvana Ramabhadran

Online Network Inference from Graph-Stationary Signals with Hidden Nodes 13456
Andrei Buciulea, Madeline Navarro, Samuel Rey, Santiago Segarra, Antonio G. Marques

Universal Training of Neural Networks to Achieve Bayes Optimal Classification Accuracy 13461
Mohammadreza Tavasoli Naeini, Ali Bereyhi, Morteza Noshad, Ben Liang, Alfred O. Hero

GraFPrint: A GNN-Based Approach for Audio Identification 13466
Aditya Bhattacharjee, Shubhr Singh, Emmanouil Benetos

Earbuds Orientation Alignment Based on Markov Chain Monte Carlo Sampling 13471
Xianghao Zhan, Nafiul Rashid, Ebrahim Nemati, Mohsin Y Ahmed, Jilong Kuang

ConvexECG: Lightweight and Explainable Neural Networks for Personalized, Continuous Cardiac Monitoring 13476
Rayan Ansari, John Cao, Sabyasachi Bandyopadhyay, Sanjiv M. Narayan, Albert J. Rogers, Mert Pilanci

Towards Clinically Feasible Nonintrusive Quality and Intelligibility Indices for Hearing Aids 13481
Vahid Ashkanichenarlogh, Paula Folkeard, Vijay Parsa

Freeze and Learn: Continual Learning with Selective Freezing for Speech Deepfake Detection 13486
Davide Salvi, Viola Negroni, Luca Bondi, Paolo Bestagini, Stefano Tubaro

Knowledge Enhanced Multi-Domain Recommendations in an AI Assistant Application 13491
Elan Markowitz, Ziyan Jiang, Fan Yang, Xing Fan, Zheng Chen, Greg Ver Steeg, Aram Galstyan

Self-Supervised Multimodal Speech Representations for the Assessment of Schizophrenia Symptoms 13496
Gowtham Premananth, Carol Espy-Wilson

Semantic Hierarchical Prompt Tuning for Parameter-Efficient Fine-Tuning 13501
Haowei Zhu, Fangyuan Zhang, Rui Qin, Tianxiang Pan, Junhai Yong, Bin Wang

Heart Sounds for High Blood Pressure Prediction 13506
Erika Bondareva, Jing Han, Katarzyna Szczurek, Dawid Szczepanek, Tomasz Jadczyk, Cecilia Mascolo

Diffusion-Based Identity-Preserving Facial Privacy Protection 13511
Dong Han, Salaheldin Mohamed, Yong Li, Joachim Denzler

Reconciling AMP Algorithms Derived from Belief Propagation Or the Large System Limit Bethe Free Energy 13516
Zilu Zhao, Fangqing Xiao, Christo K. Thomas, Dirk Slock

Joint Energy-Based Optimization of Binary Offloading Decisions and Communication Resources in TDMA Systems, Via Dynamic Programming 13521
Ruilin Ji, Sorina Dumitrescu

PlantPCC: Dual Sampling and Multi-Level Geometry-Aware Contrastive Regularization for Plant Point Cloud Completion 13526
Xiaomeng Li, Wenxu Wang, Haoxiang Sun, Yanhao Ding, Zhenbo Li

No Class Left Behind: A Closer Look at Class Balancing for Audio Tagging 13531
Janek Ebbers, François G. Germain, Kevin Wilkinghoff, Gordon Wichern, Jonathan Le Roux

Noise-Agnostic Multitask Whisper Training for Reducing False Alarm Errors in Call-For-Help Detection 13536
Myeonghoon Ryu, June-Woo Kim, Minseok Oh, Suji Lee, Han Park

LLM Based Text Generation for Improved Low-Resource Speech Recognition Models 13541
Tohru Nagano, Gakuto Kurata, Samuel Thomas, Hong-Kwang J. Kuo, Daniel Bolanos, Hyun Jung, George Saon

Efficient Finetuning for Dimensional Speech Emotion Recognition in the Age of Transformers 13546
Aneesha Sampath, James Tavernor, Emily Mower Provost

Data Efficient Child-Adult Speaker Diarization with Simulated Conversations 13551
Anfeng Xu, Tiantian Feng, Helen Tager-Flusberg, Catherine Lord, Shrikanth Narayanan

Towards Zero-Shot Cross-Lingual SLU with Syntax-Aware Multi-View Contrastive Learning 13556
Yuxin Xie, Zhen Xiong, Tinghe Zhang, Mengke Cui, Yuqi Li, Zhiqi Huang, Zhihong Zhu

SBL Algorithms for the Multiple Measurement Vector Problem: New Modeling and Inference Methods 13561
Vinay Kanakeri, Florian Meyer, Bhaskar D. Rao

Efficient Fusion of Computationally Diverse Modalities Using Chunking and Cross-Attention 13566
Christian Flores, Lucas Goncalves, Carlos Busso

Exploring the Differences Between Deaf and Hearing Infant Cries 13571
Enjamamul Hoq, Ifeoma Nwogu

LiDAR Light Scattering Augmentation (LISA): Physics-Based Simulation of Adverse Weather Conditions for 3D Object Detection 13576
Velat Kilic, Deepti Hegde, A. Brinton Cooper, Vishal M. Patel, Mark Foster

PDMX: A Large-Scale Public Domain MusicXML Dataset for Symbolic Music Processing 13581
Phillip Long, Zachary Novack, Taylor Berg-Kirkpatrick, Julian McAuley

Hybrid Losses for Hierarchical Embedding Learning 13586
Haokun Tian, Stefan Lattner, Brian McFee, Charalampos Saitis

Stable Test-Time Training for Semantic Segmentation with Output Contrastive Loss 13591
Yunlong Zhang, Zhongyi Shui, Honglin Li, Yuxuan Sun, Chenglu Zhu, Lin Yang

Guess What I Think: Streamlined EEG-To-Image Generation with Latent Diffusion Models 13596
Eleonora Lopez, Luigi Sigillo, Federica Colonnese, Massimo Panella, Danilo Comminiello

Domain-Specific Adaptation in Speech Emotion Recognition Using Emotional Distribution Alignment 13601
Abinay Reddy Naini, Donita Robinson, Elizabeth Richerson, Carlos Busso

An Efficient and Streaming Audio Visual Active Speaker Detection System 13606
Arnav Kundu, Yanzi Jin, Mohammad Sekhavat, Maxwell Horton, Danny Tormoen, Devang Naik

Reducing the Sensitivity of Neural Physics Simulators to Mesh Topology Via Pretraining 13611
Nathan Vaska, Justin Goodwin, Robin Walters, Rajmonda S. Caceres

Enhancing Age-Related Robustness in Children Speaker Verification..... 13616
Vishwas M. Shetty, Jiusi Zheng, Steven M. Lulich, Abeer Alwan

An Abnormal Audio Generation Method for Fault Diagnosis of Power Transformers 13621
Ben Niu, Yangjie Wei, Ke Zhang, Zhuoran Yu

Text2FX: Harnessing CLAP Embeddings for Text-Guided Audio Effects..... 13626
Annie Chu, Patrick O'Reilly, Julia Barnett, Bryan Pardo

Unveiling Performance Bias in ASR Systems: A Study on Gender, Age, Accent, and More..... 13631
Maliha Jahan, Priyam Mazumdar, Thomas Thebaud, Mark Hasegawa-Johnson, Jesús Villalba, Najim Dehak, Laureano Moro-Velazquez

GLoG-CSUnet: Enhancing Vision Transformers with Adaptable Radiomic Features for Medical Image Segmentation 13636
Niloufar Eghbali Zarch, Hassan Bagher-Ebadian, Tuka Alhanai, Mohammad M. Ghassemi

Exponential Convergence of Stochastic Mirror Descent in Over-Parameterized Linear Models..... 13641
K Nithin Varma, Babak Hassibi

SOLVE: Spatially Optimized Lung Volume Evidence Model for Efficient Nodule Malignancy Classification 13646
Sadaf Khademi, Anastasia Oikonomou, Arash Mohammadi

Deep Learning for Modulo Sampling of FRI Signals..... 13651
Sem Koenen, Vincent Van De Schaft, Ruud J. G. van Sloun

Multi-Objective Representation Based Dynamic Prototype Learning for Unsupervised DCE-MRI Breast Tumor Segmentation 13656
Li Wang, Lihui Wang, Lei Tang, Zixiang Kuai, Jian Zhang, Hongjiang Wei

Dual-Path Contrastive Short Text Clustering with High-Order Random Walk 13661
Zhengzhong Zhu, Binjie Sun, Xuejie Zhang, Jin Wang, Xiaobing Zhou

Single-View Clothed Human Reconstruction Using Symmetric Feature 13666
Yanping Fu, Zhuangzhuang Zhao, Hao Geng, Haifeng Zhao

Multi-Scale Denoising in the Feature Space for Low-Light Instance Segmentation 13671
Joanne Lin, Nantheera Anantrasirichai, David Bull

ID-RWKV: Image Deraining RWKV 13676
Yong Yang, Jiaxuan Yang, Shuying Huang

Impact of Temporal Precision on Speech Synthesis Accuracy from Electrocorticographic Brain Signals 13681
Qinwan Rabbani, Matthew S. Fifer, Nathan E. Crone, Laureano Moro-Velazquez

Collaborative Automotive Radar Sensing Via Mixed-Precision Distributed Array Completion 13686
Arian Eamaz, Farhang Yeganegi, Yunqiao Hu, Mojtaba Soltanalian, Shunqiao Sun

Self-Supervised Prosody Learning at Phoneme-Level with Momentum Contrast for Speech Synthesis 13691
Zhao-Ci Liu, Ya-Jun Hu, Liping Chen, Zhen-Hua Ling

Fusion of Information in Multiple Particle Filtering in the Presence of Unknown Static Parameters 13696
Xiaokun Zhao, Marija Iloska, Yousef El-Laham, Monica F. Bugallo

Audiopedia: Audio QA with Knowledge 13701
Abhirama Subramanyam Penamakuri, Kiran Chhatre, Akshat Jain

Streamlining UNO: A Generalized Sampling Approach to Optimal One-Bit Modulo Sensing 13706
Arian Eamaz, Farhang Yeganegi, Mojtaba Soltanalian

Star Operation in Self-Attention for 3D Human Pose Estimation 13711
Hao Wang, Xiaochuan Wang, Ruijun Liu, Xiaoming Chen, Haisheng Li

SABiT-MNet: Scale-Adaptive Autoencoder with BiT-M Model for Identifying AMD Grades 13716
Niveen Nasr El-Den, Mohamed Elsharkawy, Mohammed Ghazal, Ali Mahmoud, Harpal Sandhu, Hani Mahdi, Ayman El-Baz

Latent Representation Learning for Multimodal Brain Activity Translation 13721
Arman Afrasiyabi, Dhananjay Bhaskar, Erica L. Busch, Laurent Caplette, Rahul Singh, Guillaume Lajoie, Nicholas B. Turk-Browne, Smita Krishnaswamy

Video-Poetry Retrieval with Multimodal Knowledge Graph Guided Unsupervised Pre-Training 13726
Xinru Wei, Yuqing Li, Bin Wu

L2 · M = C2 Large Language Models Are Covert Channels 13731
Simen Gaure, Stefanos Koffas, Stjepan Picek, Sondre Rønjom

What Does an Audio Deepfake Detector Focus On? a Study in the Time Domain 13736
Petr Grinberg, Ankur Kumar, Surya Koppisetti, Gaurav Bharaj

Multi-Style Facial Sketch Synthesis Through Masked Generative Modeling 13741
Bowen Sun, Guo Lu, Shibao Zheng

Incorporating Improved Sinusoidal Threshold-Based Semi-Supervised Method and Diffusion Models for Osteoporosis Diagnosis 13746
Wenchi Ke, Hu Chen, Yi Zhang, Xiong Deng

Identity-Preserving Audio-Driven Holistic Human Motion Video Generation 13751
Haiwei Xue, Zhensong Zhang, Minglei Li, Zonghong Dai, Zhiyong Wu

JointSwinUNETR: An Efficient Feature-Enhanced Architecture for Small Intestine Cine MRI Segmentation 13756
Yue Wang, Wenhui Li, Ziming Wang, Taoli Du, Ming Ma, Mengchao Zhang

Multi-Prototype Grouping for Continual Learning in Visual Question Answering 13761
Licheng Zhang, Zhendong Mao, Yixing Peng, Zheren Fu, Yongdong Zhang

Data-Driven Mispronunciation Pattern Discovery for Robust Speech Recognition 13766
Anna Seo Gyeong Choi, Jonghyeon Park, Myungwoo Oh

Conditional Convolutions for End-To-End Single-Stage Video Text Detection 13771
Xiaoge Song, Danhuai Zhao, Wei Zhu, Kang Zheng, Tong Lu

Source-Aware Spatial Self-Supervision for Sound Event Localization and Detection 13776
Yuto Nozaki, Yoshiaki Bando, Masaki Onishi

LEP: Leveraging Local Entropy Pruning for Sparsity in Large Language Models 13781
Yuli Chen, Bo Cheng, Yingying Zhang, Shuhao Zhang, Zhixuan Wu, Fanshen Meng

Subjective Fidelity Assessment of Audio- And Video-Driven Talking Head Generation Methods 13786
Anthony Trioux, Yusong Gao, Jiarun Song, Wenjie Wu, Faming Ma, Fuzheng Yang

Jack of All Trades, Master of None: PMP-Guided Adaptive Multi-Teacher Distillation with Meta-Learning 13791
Sisi Zhang, Zechao Lin, Xingbin Wang, Yulan Su, Yan Wang, Rui Hou, Dan Meng

Resource-Efficient and Noise-Robust Modality Fusion for Audio-Visual Speech Recognition 13796
Alexandros Koumparoulis, Gerasimos Potamianos

META-CAT: Speaker-Informed Speech Embeddings Via Meta Information Concatenation for Multi-Talker ASR 13801
Jinhan Wang, Weiqing Wang, Kunal Dhawan, Taejin Park, Myungjong Kim, Ivan Medennikov, He Huang, Nithin Koluguri, Jagadeesh Balam, Boris Ginsburg

Dynamic SpikFormer: Low-Latency & Energy-Efficient Spiking Neural Networks with Dynamic Time Steps for Vision Transformers 13806
Gourav Datta, Zeyu Liu, Anni Li, Peter A. Beerel

Variance-Reduced Clipping for Non-Convex Optimization 13811
Amirhossein Reisizadeh, Haochuan Li, Subhro Das, Ali Jadbabaie

Asymptotic Behavior Analysis of Antenna Selection Via Sparsity-Induced Precoder 13816
Xiuxiu Ma, Abla Kammoun, Mohamed-Slim Alouini, Tareq Y. Al-Naffouri

SSM2Mel: State Space Model to Reconstruct Mel Spectrogram from the EEG 13821
Cunhang Fan, Sheng Zhang, Jingjing Zhang, Zexu Pan, Zhao Lv

Keeping Your Eyes on the Fingertip: A Two-Stage In-Air Handwritten Recognition Method 13826
Zeyu Qiu, Weiqiang Wang

Multi-Modal Speech Enhancement with Limited Electromyography Channels 13831
Fuyuan Feng, Longting Xu, Rohan Kumar Das

Prompt-Augmented Feature with Cross-Domain Contrastive Learning for Efficient Multi-Domain Sentiment Analysis 13836
Rui Li, Tao Yang, Cheng Liu, Dazhi Jiang, Hau-San Wong, Si Wu

Alignment-Free Training for Transducer-Based Multi-Talker ASR 13841
Takafumi Moriya, Shota Horiguchi, Marc Delcroix, Ryo Masumura, Takanori Ashihara, Hiroshi Sato, Kohei Matsuura, Masato Mimura

DAREK - Distance Aware Error for Kolmogorov Networks 13846
Masoud Ataei, Mohammad Javad Khojasteh, Vikas Dhiman

Intra-Modal Relation and Emotional Incongruity Learning Using Graph Attention Networks for Multimodal Sarcasm Detection 13851
Devraj Raghuvanshi, Xiyuan Gao, Zhu Li, Shubhi Bansal, Matt Coler, Nagendra Kumar, Shekhar Nayak

Exploiting Application-To-Architecture Dependencies for Designing Scalable OS 13856
Yao Xiao, Nikos Kanakaris, Anzhe Cheng, Chenzhong Yin, Nesreen K. Ahmed, Shahin Nazarian, Andrei Irimia, Paul Bogdan

Improved Recognition of the Speech of People with Parkinson's Who Stutter..... 13861
Jonghwan Na, Xiuwen Zheng, Bowon Lee, Mark Hasegawa-Johnson

Identifying and Mitigating Mismatched Language Code in Multilingual ASR..... 13866
Jaeyoung Kim, Sepand Mavandadi, Kartik Audhkhasi, Shikhar Bharadwaj, Brian Farris, Tongzhou Chen, Bhuvana Ramabhadran, Sriram Ganapathy

CGEDN: Approximation of Graph Edit Distance with Path Generation Via Learning Node Matching..... 13871
Liu Yang, Qiankun Zheng, Zidong Wang

UniIVFT: Towards a Unified Framework for Infrared-Visible Fusion and Translation 13876
Honglin Wu, Xueqiong Li, Shaowu Yang, Huibin Tan, Yuhua Tang, Tianrui Liu

M-MoE: Mixture of Mixture-Of-Expert Model for CTC-Based Streaming Multilingual ASR 13881
Songjun Cao, Xiong Wang, Yike Zhang, Xiaoming Zhang, Long Ma

Improving Lip-Synchrony in Direct Audio-Visual Speech-To-Speech Translation 13886
Lucas Goncalves, Prashant Mathur, Xing Niu, Chandrashekhar Lavania, Brady Houston, Srikanth Vishnubhotla, Lijia Sun, Anthony Ferritto

Hierarchical Spatiotemporal Attention Network for Fine-Grained Brain Cognitive State Recognition 13891
Yike Wu, Ning An, Zixuan Zeng, YouYong Kong

SoundTRC: DNN-Based Acoustic Target Region Control..... 13896
Yuhang He, Andrew Markham, Okan Köpüklü

Clinically Robust Polyp Segmentation: Enhanced Generalization and Perturbation Resistance 13901
Shanchuan Wang, Fang Yuan, Jian-Nan Su, Min Gan

Fast DPCNs for Feature Extraction Without Labels..... 13906
Wenqian Xue, Chi Ding, Jose C. Principe

From Pixels to Voice: A Simple and Efficient End-To-End Spoken Image Description Approach Via Vision Codec Language Models.....13911
Chung Tran, Sakriani Sakti

Can Automated Speech Recognition Errors Provide Valuable Clues for Alzheimer's Disease Detection? 13916
Yin-Long Liu, Rui Feng, Ye-Xin Lu, Jia-Xin Chen, Yang Ai, Jia-Hong Yuan, Zhen-Hua Ling

Investigating F0 Estimation in Speech Synthesis from Real-Time MRI Articulatory Data 13921
Yuto Otani, Shun Sawada, Hidefumi Ohmura, Kouichi Katsurada

SPNet: Sparse-Mask Prompt-Learning Network for Cerebrovascular Segmentation 13926
Wenqi Shan, Qiang Li, Zhiwei Wang

Decentralized Online Ensembles of Gaussian Processes for Multi-Agent Systems 13931
Fernando Llorente, Daniel Waxman, Petar M. Djurić

Enhancing Speech Emotion Recognition with Speech Dynamic Modeling and Multi-Modal Knowledge Distillation 13936
Chuanbo Zhu, Chao Sun, Yifan Liu, Jincai Chen, Ke Luo

Artistic Image Aesthetics Assessment Assisted by Photographic Visual Attributes 13941
Haiyong Tang, Yihua Chen, Xiaoping Liang, Lv Chen, Pengsheng Huang, Zhenjun Tang

Two Heads Are Better than One: Averaging Along Fine-Tuning to Improve Targeted Transferability 13946
Hui Zeng, Sanshuai Cui, Biwei Chen, Anjie Peng

Cohort-Sensitive Labeling: An Effective Approach for Enhancing ASR Performance 13951
Jonghwan Na, Mark Hasegawa-Johnson, Bowon Lee

TCTformer: Long-Term Forecasting with Dual Attention Transformers 13956
Long Sun, Xiaoyan Gao, Kai Xia, Xuwei Hu, Yuan Feng

MINR: Efficient Implicit Neural Representations for Multi-Image Encoding 13961
Wenyong Zhou, Taiqiang Wu, Zhengwu Liu, Yuxin Cheng, Chen Zhang, Ngai Wong

SSRMamba: Efficient Visual State Space Model for Spectral Super-Resolution 13966
Baisong Li, Xingwang Wang, Haixiao Xu

HAPG-SAQAM: Human Auditory Perception Guided Spatial Audio Quality Assessment Metric 13971
Yuanming Zheng, Jiaxuan Yao, Xiangyu Deng, Yuhong Yang, Ruiqi Liao, Weiping Tu, Cedar Lin

Efficient Defocus Deblurring Networks Based on Diffusion Models 13976
Kang Chen, Yuanjie Liu

A Metric for Predicting the Quality of Ambisonic Spatial Audio Reproduced Using Spatially Interpolated Or Extrapolated Room Impulse Responses 13981
Hualin Ren, Christian Ritz, Jiahong Zhao, Xiguang Zheng, Daeyoung Jang

Multi-Scale Conditional Generative Adversarial Networks for Wind Speed Data Imputation in Earthen Ruins Protection 13986
Hang Li, Hai Wang, Rui Cao, Shuo Ji, Jie Zheng

Under-Counted Matrix Completion Without Detection Features 13991
Tri Nguyen, Shahana Ibrahim, Rebecca A. Hutchinson, Xiao Fu

KABON: Knowledge Aggregation with Vision-Language Model for Black-Box Open-Set Domain Adaptation 13996
Zhixin Zeng, Yusen Zhang, Ji Wang

Revisiting and Refining Lagunas' Beamforming for Acoustic Imaging 14001
Xun Wang, Jérôme Antoni, Jianing Li, Jean-Daniel Chazot, Jing Lin

DrLLM: Prompt-Enhanced Distributed Denial-Of-Service Resistance Method with Large Language Models 14006
Zhenyu Yin, Shang Liu, Guangyuan Xu

Advanced Graph-MLPs Distillation Based on Global and Local Hyperbolic Geometry Learning 14011
Yao Guan, Wenzhu Yan, Ting Yuan, Yanmeng Li

Contactless Nighttime Stress Monitoring with mmWave Radar 14016
Xiaohan Xu, Dongheng Zhang, Zhi Lu, Jinbo Chen, Zhi Wu, Ruixu Geng, Qibin Sun, Yan Chen

Collaborative Personalized Federated Learning Via Exponential Moving Average Optimization 14021
Yuqing Li, Jintao Liang, Peng Tang, Sen Su

CascadePAIE: Reallocating Relevance for Event Roles and Event Text in Event Argument Extraction 14026
Chunyu Yao, Yi Guo

Personalized Federated Class-Incremental Learning Through Critical Parameter Transfer 14031
Feng Wu, Siwei Feng, Yuanlu Chen, Libang Zhao

Speech Emotion Recognition Based on Large-Scale Automatic Speech Recognizer 14036
Ryo Fukuda, Takatomo Kano, Atsushi Ando, Atunori Ogawa

Error Bounds Revisited, and How to Use Bayesian Statistics While Remaining a Frequentist 14041
Ning Xu, Christopher M. Foster, Jonathan H. Manton

Regularized Domain Adaptation for Estimation Tasks in Partially Observed Target Domains 14046
Varun Kelkar, H.S. Melihcan Erol, Muhammad Aneeq Uz Zaman, Omer Tanovic, Ravi Kiran Raman

Audio Sparse-Transformer for Speech Classification 14051
Hassan Salami Kavaki, Michael I Mandel

Enhancing Robustness of Implicit Neural Representations Against Weight Perturbations 14056
Wenyong Zhou, Yuxin Cheng, Zhengwu Liu, Taiqiang Wu, Chen Zhang, Ngai Wong

Martin: Mobility-Aware Reputation Mechanism for Federated Learning 14061
Lixin Liu, Xin Chang, Hanqing Yang, Jingyu Wang, Xiaolin Zhang, Cuiyun Shi

SCF-Stega: Controllable Linguistic Steganography Based on Semantic Communications Framework 14066
Yilin Long, Zhongliang Yang, Zhuang Wang, Zhili Zhou, Yongfeng Huang, Linna Zhou

GATOmics: A Novel Multi-Omics Graph Attention Network Model for Cancer Driver Gene Detection 14071
Ge Kong, Jiao Wang, Juan Wang

TriFP-NGram: Integrating Three Complementary Fingerprint and N-Gram Features for Enhanced Drug-Target Affinity Prediction 14076
Jiao Wang, Ge Kong, Juan Wang

Adversarial Training and Cross-Modal Feature Fusion in Multimodal Sentiment Analysis 14081
Junhuai Li, Chuang Lin, Huaijun Wang, Yuxing Zhi, Jing Chen, Tao Huang

Dual Path Unsupervised Real Image Denoising 14086
Yao Li, Zheng-jun Zha

Enhancing Federated Knowledge Distillation in Heterogeneous and Non-IID Scenarios 14091
Wenjie Lv, Yu He, Sen Liu, Xingjun Ma, Xiang Liu, Guangnan Ye, Hongfeng Chai

Super Capacity SRS Design for 5G and Beyond Using Channel In-Painting 14096
Usman Akram, Fan Zhang, Shawn Ma, Yang Li, Haris Vikalo

VAGeo: View-Specific Attention for Cross-View Object Geo-Localization 14101
Zhongyang Li, Xin Yuan, Wei Liu, Xin Xu

Sketch-Based Point Cloud Generation with Diffusion Model and Pre-Training Enhancement 14106
Yangdong Chen, Mohan Chen, Yuejie Zhang, Rui Feng, Tao Zhang, Shang Gao

Text-Guided Few-Shot Semantic Segmentation with Training-Free Multimodal Feature Matching 14111
Guillaume Buthmann, Tomoya Sakai, Haoxiang Qiu, Takayuki Katsuki, Daiki Kimura

ATP-TTS: Adaptive Thresholding Pseudo-Labeling for Low-Resource Multi-Speaker Text-To-Speech 14116
Feng Li, Shen Chen, Hanjin Yang, Shupei Yuan

Event-Based Video Person Re-Identification Via Cross-Modality and Temporal Collaboration 14121
Renkai Li, Xin Yuan, Wei Liu, Xin Xu

Enhancing Information Extraction with METORIE: A Metaphor and Trap-Based Dataset for Cross-Domain Fine-Tuning 14126
Zhengyuan Pan, Yilian Peng, Zhongquan Jian, Yanhao Chen, Wentao Qiu, Haonan Ma, Junfeng Yao, Meihong Wang, Qingqiang Wu

Leveraging Chain of Thought Towards Empathetic Spoken Dialogue Without Corresponding Question-Answering Data 14131
Jingran Xie, Shun Lei, Yue Yu, Yang Xiang, Hui Wang, Xixin Wu, Zhiyong Wu

A Diffusion Model Over Directed Acyclic Graphs for Event Schema Generation 14136
Guoxuan Ding, Haotian Jin, Xiaobo Guo, Xin Wang, Nan Mu, Lei Wang, Daren Zha

Low-Rank Adaptation Method for Respiratory Sound Classification: A Necessary Road Towards Large Models 14141
Gaoyang Dong, Yufei Shen, Jianhong Wang, Shunwang Xie, Minghui Zhang, Ping Sun

Dynamic Dictionary Design for Localization in Automotive Radar Systems 14146
Farhan Bishe, Mohammed Saif, Jun Li, Shahrokh Valaee

Improved Extrinsic Calibration of Acoustic Cameras Via Batch Optimization 14151
Zhi Li, Jiang Wang, Xiaoyang Li, He Kong

MP-DPCC: A Motion Proxy-Based Dynamic Point Cloud Compression Framework 14156
ZhaoYi Jiang, Dong Han, Cao Song, FangZhe Nan, Bailin Yang

FUTGA-MIR: Enhancing Fine-Grained and Temporally-Aware Music Understanding with Music Information Retrieval 14161
Junda Wu, Zachary Novack, Amit Namburi, Hao-Wen Dong, Carol Chen, Jiaheng Dai, Julian McAuley

Leave No Stone Unturned: Optimizing Subpattern Information Entropy for Coreset Selection 14166
Haohao Song, Qiao Xiang, Jiwu Shu

Multi-Feature Audio Fusion for Nonverbal Vocalization Classification 14171
Siddhant Bikram Shah, Kristina T. Johnson

Robust Semantic Communications for Speech Transmission 14176
Zhenzi Weng, Zhijin Qin, Geoffrey Ye Li

Find Details in Long Videos: Tower-Of-Thoughts and Self-Retrieval Augmented Generation for Video Understanding 14181
Tong Yue, Mingrui Xiao, Dafeng Zhang, Xin Liu, Yali Li, Shengjin Wang

Self-Trained Model for ECG Complex Delineation 14186
Aram Avetisyan, Nikolas Khachaturov, Ariana Asatryan, Shahane Tigranyan, Yury Markin

DistillW2N: A Lightweight One-Shot Whisper to Normal Voice Conversion Model Using Distillation of Self-Supervised Features 14191
Tianyi Tan, Haoxin Ruan, Xinan Chen, Kai Chen, Zhibin Lin, Jing Lu

Causal Speech Enhancement Based on a Two-Branch Nested U-Net Architecture Using Self-Supervised Speech Embeddings 14196
Seorim Hwang, Sung Wook Park, Youngcheol Park

DX2CT: Diffusion Model for 3D CT Reconstruction from Bi Or Mono-Planar 2D X-Ray(s) 14201
Yun Su Jeong, Hye Bin Yoo, Il Yong Chun

Query-Aware Temporal Aggregation Network for Temporal Knowledge Graph Reasoning 14206
Shihao Liu, Xiaofei Zhou

MTMDC-GAN: Self-Attention Driven Multi-Scale Temporal Synthesis with Multi-Domain Analysis and Contrastive Learning 14211
Weihai Zhi, Kejing He

A Joint Time-Frequency Attention for Leakage Detection in Water Distribution Networks Using Time Series Decomposition 14216
Juan Luo, Yiyang Chen, Jielong Yang, Xionghu Zhong

Knowledge-Enhanced Poetry-Image Synthesis with Large Language Model 14221
Simin Yang, Yuqing Li, Bin Wu

FR2ViT: Finetuning-Free Token Reduction for Dense Prediction Through a Refinement-Reactivation Architecture 14226
Haipeng Fang, Ziheng Wu, Xinyi Zou, Jun Huang, Juan Cao, Sheng Tang

Denoising and Restoring Channel State Information for 5G Indoor Positioning in Low-SNR Scenarios 14231
Jianing Chen, Junze Yang, Chuhao Chen, Xiangxu Meng, Wenqi Zheng

Multi-Level Encoder with Global Topic for Task-Oriented Dialogue Summarization 14236
Zhuoqi He, Peijie Huang, Yuhong Xu, Youming Peng, Mingzhi Xu, Xinyang Lin

The Impact of Deployment Height on the Mean Delay in Large UAV Networks 14241
S. Tayyaba, S. Vanka

Trimformer: A Novel Sequence Compression Mechanism with Local Attention 14246
Ran Dou, Liyang Ru, Jose Principe

Enhancing Privacy in Radar-Based Vital Sign Monitoring Via Non-Linear FMCW Waveforms 14251
Zhihao Tao, Athina Petropulu

Improving Adversarial Transferability Through Channel-Wise Scaling and Frequency-Random Dropping 14256
Pei Chen, Zhiyong Feng, Meng Xing, Yiming Zhang, Jinqing Zheng

CoLLAP: Contrastive Long-Form Language-Audio Pretraining with Musical Temporal Structure Augmentation 14261
Junda Wu, Warren Li, Zachary Novack, Amit Namburi, Carol Chen, Julian McAuley

Can Fairness and Robustness Be Simultaneously Achieved Under Byzantine Attacks? 14266
Huigan Zheng, Runhua Wang, Xiao Wang, Qing Ling

Can Quality Survive Scale? Toward an Equal-Quality Instance-Dependent Label Noise Model 14271
Haohao Song, Qiao Xiang, Jiwu Shu

InfoMin-Based Query Embedding Optimization for Query-Based Universal Sound Separation 14276
Zhen Wang, Jiqing Han, Liwen Zhang, Youcheng Zhang

EPCPE: A Real-Time End-To-End Pipeline for RGB-Based Category-Level 6D Pose Estimation 14281
Xiaofeng Fan, Jie Guo, Shichao Kan, Yixiong Liang

Fast Sparse Learning from Streaming Data with LASSO 14286
Marija Iloska, Petar M. Djurić, Mónica F. Bugallo

VisTa: Visual-Contextual and Text-Augmented Zero-Shot Object-Level OOD Detection 14291
Bin Zhang, Xiaoyang Qu, Guokuan Li, Jiguang Wan, Jianzong Wang

DiffCSS: Diverse and Expressive Conversational Speech Synthesis with Diffusion Models 14296
Weihao Wu, Zhiwei Lin, Yixuan Zhou, Jingbei Li, Rui Niu, Qinghua Wu, Songjun Cao, Long Ma, Zhiyong Wu

Explainable Adversarial Attacks on Coarse-To-Fine Classifiers 14301
Akram Heidarizadeh, Connor Hatfield, Lorenzo Lazzarotto, HanQin Cai, George Atia

ConPCO: Preserving Phoneme Characteristics for Automatic Pronunciation Assessment Leveraging Contrastive Ordinal Regularization 14306
Bi-Cheng Yan, Yi-Cheng Wang, Jiun-Ting Li, Meng-Shin Lin, Hsin-Wei Wang, Wei-Cheng Chao, Berlin Chen

Low Rank and Sparse Fourier Structure in Recurrent Networks Trained on Modular Addition 14311
Akshay Rangamani

Time-Space-Interlaced Spatiotemporal Graph Forecasting Via Two-Stage Summarized Attention 14316
Zhaoyang Sun, Yudong Zhang, Xuan Yu, Kai Wang, Binwu Wang, Yang Wang, Xu Wang

Hyperspectral Image Reconstruction with Unseen Material Detection 14321
Ye Ma, Songnan Lin, Bihan Wen

Contextualization of ASR with LLM Using Phonetic Retrieval-Based Augmentation 14326
Zhihong Lei, Xingyu Na, Mingbin Xu, Ernest Pusateri, Christophe Van Gysel, Yuanyuan Zhang, Shiyi Han, Zhen Huang

Bottleneck-Constrained Contrastive Decoupled Network for Multimodal Aspect-Based Sentiment Classification 14331
Rui Liu, Jiahao Cao, Lei Jiang, Chaodong Tong, Haimei Qin, Yanan Cao

Singing Voice Conversion with Accompaniment Using Self-Supervised Representation-Based Melody Features 14336
Wei Chen, Binzhu Sha, Jing Yang, Zhuo Wang, Fan Fan, Zhiyong Wu

Inter-Frame Skip Coding Mode for Point Cloud Geometry Compression in Solid G-PCC 14341
Ren Huang, Wei Zhang, Jiangxue Han, Fuzheng Yang

Small Target Insect Detection Based on Improved YOLOv8n 14346
Jianyu Shi, Yuan Jia, Gang Zhou, Jiajia Wang, Zhenhong Jia

Quaternion CNN with Salient Features for Color Image Denoising 14351
Yi Liu, Qiyu Jin, Jie Yang

Hybrid Offline Passive Grammatical Inference and Online Planning for Non-Markovian Tasks 14356
Mahyar Alinejad, Alvaro Velasquez, Yue Wang, George Atia

PGDGS: Improving Few-Shot 3D Gaussian Splatting with Progressive Gaussian Densification 14361
Haoyang Huang, Zhe Zhang, Guanhua Wu, Ronggang Wang

Bridging Speech and Text Foundation Models with ReShape Attention 14366
Takatomo Kano, Atsunori Ogawa, Marc Delcroix, William Chen, Ryo Fukuda, Kohei Matsuura, Takanori Ashihara, Shinji Watanabe

Causality-Guided Context-Aware Multimodal Public Speaking Anxiety Detection for Out-Of-Distribution Generalization 14371
Tingting Zhang, Jiachen Tan, Zihua Xiong, Bin Wu, Chunping Zheng

HazeCLIP: Towards Language Guided Real-World Image Dehazing 14376
Ruiyi Wang, Wenhao Li, Xiaohong Liu, Chunyi Li, Zicheng Zhang, Xiongkuo Min, Guangtao Zhai

PointActionCLIP: Preventing Transfer Degradation in Point Cloud Action Recognition with a Triple-Path CLIP 14381
Wei Tao, Shenglin He, Xiaoyang Qu, Jiguang Wan, Jianzong Wang

Dynamically Causal-Enhanced Exercise Representations for Adaptive Knowledge Tracing 14386
Yanhong Bai, Jiabao Zhao, Tingjiang Wei, Jinxin Shi, Liang He

Peer-To-Peer Learning Dynamics of Wide Neural Networks 14391
Shreyas Chaudhari, Srinivasa Pranav, Emile Anand, José M. F. Moura

Personalized Graph Transformer for Federated Graph Learning 14396
Haohe Jia, Yi Huang, Hongbin Zhu, Hongfeng Chai

Generative Expansion of Small Datasets: An Expansive Graph Approach 14401
Vahid Jebraeeli, Bo Jiang, Hamid Krim, Derya Cansever

Integrating Audio Narrations to Strengthen Domain Generalization in Multimodal First-Person Action Recognition 14406
Cagri Gungor, Adriana Kovashka

Decoding the Unintelligible: Neural Speech Tracking in Low Signal-To-Noise Ratios 14411
Xiaomin He, Vinay S Raghavan, Nima Mesgarani

CrossSleep: Multi-Scale Attention with Cross-Time Learning for Single Channel EEG-Based Sleep Staging 14416
Jingchuan Lu, Jinlong Yang

MAITFuse:Multi-Dimension Adaptive Interaction Transform Network for Infrared-Visible Image Fusion 14420
Yabin Sun, Wentai Lei, Ziyi Zhang, Jiongchang Liu, Chenxu Li, Tao Zhang

A Comparative Study of Invariance-Aware Loss Functions for Deep Learning-Based Gridless Direction-Of-Arrival Estimation 14425
Kuan-Lin Chen, Bhaskar D. Rao

Compact Neural TTS Voices for Accessibility 14430
Kunal Jain, Eoin Murphy, Deepanshu Gupta, Jonathan Dyke, Saumya Shah, Vasilieios Tsiaras, Petko Petkov, Alistair Conkie

NanoGen: A High-Affinity Nanobody Generation Model with Guided Diffusion 14435
Dezhi Wu, Xuejiao Liu, Yiming Qin, Stephanie M. Linker, Karin Hrovatin, Alexander V. Hopp, Feng Tan

Learning Diffusion Model from Noisy Measurement Using Principled Expectation-Maximization Method 14440
Weimin Bai, Weiheng Tang, Enze Ye, Siyi Chen, Wenzheng Chen, He Sun

A Bandwidth Efficient Dual Function Radar Communication System Based on a MIMO Radar Using OTFS Waveforms 14445
Kailong Wang, Athina Petropulu

FedFLD: Heterogeneous Federated Learning Via Forget-Less Distillation 14450
Xiaoyang Yi, Jian Zhang, Jing Chen, Yuru Bao, Lingkai Xing

Optimum Power-Subcarrier Allocation and Time-Sharing in Multicarrier NOMA Uplink 14455
Sagnik Bhattacharya, Kamyar Rajabalifardi, Muhammad Ahmed Mohsin, John M. Cioffi

Exploring Inter-Variate and Long-Term Dependencies to Boost Multivariate Time Series Forecasting 14460
Xi Ding, Yifan He, Shuigeng Zhou, Guiyang Liu, Qi Zhou

Colored Point Cloud-Based Mesh Registration for Enhancing Inter-Frame Coding of Texture Video in V-DMC 14465
Jianxiang Sun, Wenjie Zou, Jiapan Zhao, Fuzheng Yang

DiGradPatch: Black-Box Patch Attacks Via Diffusion-Based Double Gradient and Sensitive Distribution Guidance 14470
Yang Wu, Jing Liu

Variable Bitrate Residual Vector Quantization for Audio Coding 14475
Yunkee Chae, Woosung Choi, Yuhta Takida, Junghyun Koo, Yukara Ikemiya, Zhi Zhong, Kin Wai Cheuk, Marco A. Martínez-Ramírez, Kyogu Lee, Wei-Hsiang Liao, Yuki Mitsufuji

A Multi-Modal Approach to Dysarthria Detection and Severity Assessment Using Speech and Text Information 14480
Anuprabha M, Krishna Gurugubelli, Kesavaraj V, Anil Kumar Vuppala

Enhancing Boundary-Handling Strategies for Convolutional Sparse Representation Model with 46 × 46 Convolution-Multiplication Properties 14485
Yuto Tsukiashi, Yoshimitsu Kuroki

Successive Interference Cancellation-Aided Diffusion Models for Joint Channel Estimation and Data Detection in Low Rank Channel Scenarios 14490
Sagnik Bhattacharya, Muhammad Ahmed Mohsin, Kamyar Rajabalifardi, John M. Cioffi

Improving Robustness of Post-Hoc Calibration Against Common Corruptions by Learnable Augmentation 14495
Jun Zhang, Minghao Hu, Zhunchen Luo, Wei Luo, Guotong Geng, Xiaoyin Bai

A Domain-Specific Multilingual Speech Translation Corpus Via Simultaneous Interpretation 14500
Seunghee Han, Gary Geunbae Lee, Hung Soon Kim, Sunhee Kim, Minhwa Chung

Uncertainty Estimation for Out-Of-Distribution Detection of Whole Slide Images 14505
Saba Heidari Gheshlaghi, Nasim Yahya Soltani, Masoud Ganji

Boosting the Transferability of Adversarial Examples Via Local Mixup and Adaptive Step Size 14510
Junlin Liu, Chenyu Zhang, Xinchen Lyu

Stable and Lightweight Deep Primal-Dual Unrolling for Constrained Image Restoration with Convolutional Sparse Coding 14515
Takafumi Ueki, Kazuki Naganuma, Shunsuke Ono

Neighborhood Attention Transformer with Progressive Channel Fusion for Speaker Verification 14520
Nian Li, Jianguo Wei

TROI: Cross-Subject Pretraining with Sparse Voxel Selection for Enhanced fMRI Visual Decoding 14525
Ziyu Wang, Tengyu Pan, Zhenyu Li, Ji Wu, Xiuxing Li, Jianyong Wang

TIRPL: Tailored-Made Inverse Rendering for Point-Light Scenes 14530
Zonglin Tian, Sicong Cheng, Junli Zhao, Fuqing Duan

TextHair3D: Text-Driven 3D Hair Editing with Generative Priors 14535
Xiaoxue Li, Huilong Pi, Yunchuan Qin, Ruihui Li, Kenli Li

A Novel Multimodal Method for Decoding Speech Perception from Brain Activities 14540
Aoke Zhang, Bo Wang, Xihong Wu, Jing Chen

Dual-Domain Feature-Guided Task Alignment for Enhanced Small Object Detection 14545
Fangrui Guo, Junwei Wu, Quan Zhang

A CT-Based Prediction System for Determining Respiratory Support Level in COVID-19 Patients 14550
Ahmed Sharafeldeen, Hossam Magdy Balaha, Ibrahim Shawky Farahat, Mohammed Ghazal, James Connelly, Eric van Bogaert, Ayman El-Baz

Distributionally Robust Kalman Filtering Over an Infinite-Horizon 14555
Joudi Hajar, Taylan Kargin, Vikrant Malik, Babak Hassibi

Diffusion-Based Data Augmentation for Object Counting Problems 14560
Zhen Wang, Yuelei Li, Jia Wan, Nuno Vasconcelos

Neural Ambisonic Encoding for Multi-Speaker Scenarios Using a Circular Microphone Array 14565
Yue Qiao, Vinay Kothapally, Meng Yu, Dong Yu

UCIL: An Unsupervised Class Incremental Learning Approach for Sound Event Detection 14570
Yang Xiao, Rohan Kumar Das

Locally Correctable Lattices 14575
Haodong Yang, Venkata Gandikota

DULRTC-RME: A Deep Unrolled Low-Rank Tensor Completion Network for Radio Map Estimation 14580
Yao Wang, Xin Wu, Lianming Xu, Na Liu, Li Wang

Make Some Noise: Towards LLM Audio Reasoning and Generation Using Sound Tokens 14585
Shivam Mehta, Nebojsa Jojic, Hannes Gamper

GPT-LAD: Leveraging Large Multimodal Models for Logical Anomaly Detection 14590
Yoojin An, Dongyeon Kang

A Multi-Label EEG Dataset for Mental Attention State Classification in Online Learning 14595
Huan Liu, Yuzhe Zhang, Guanjian Liu, Xinxin Du, Haochong Wang, Dalin Zhang

Joint Feature and Kernel Fusion for Improved Depth-Aware Panoptic Segmentation 14600
Yulong Bai, Shu Tian, Xin Zhao, Xu-Cheng Yin

A Margin-Maximizing Fine-Grained Ensemble Method 14605
Jinghui Yuan, Hao Chen, Renwei Luo, Feiping Nie

Enhancing Zero-Shot Emotional Voice Conversion Via Speaker Adaptation and Duration Prediction 14610
Shiyan Wang, Tianhua Qi, Cheng Lu, Zhaojie Luo, Wenming Zheng

Enhanced Breast Cancer Molecular Biomarker Classification: A Novel Two-Stage Machine Learning Pipeline for Accurate Histological Analysis of Whole Slide Images 14615
A. Aboudessouki, Kh. M. Ali, A. Alksas, M. Elsharkawy, M. ABO Rahma, M. Ghazal, N. Mekky, E. El-Daydamony, D. Gondim, A. El-Baz

Multi-Context Temporal Consistent Modeling for Referring Video Object Segmentation 14620
Sun-Hyuk Choi, Hayoung Jo, Seong-Whan Lee

Interactive and Balanced Multimodal Learning Via Cross Attention and Gradient Modulation for Compressed Video Action Recognition 14625
Xinqi Li, Shaojie Li, Ming Ma

Translational Generative Retrieval Via Potential Query Generation 14630
Yihan Guo, Tingwen Liu, Jiawei Sheng, Duohe Ma, Ming Sun, Ling Tian

Bernoulli-Gaussian Scale Mixture Model and BP Method for Multi-Snapshot Sparse Signal Recovery 14635
Shaoxiu Wei, Mingchao Liang, Bhaskar Rao, Florian Meyer

Data-Aided Regularization of Direct-Estimate Combiner in Distributed MIMO Systems 14640
Bikshapathi Gouda, Italo Atzeni, Antti Tölli

A Structured Neural Network Approach for Learning Improved Iterative Algorithms for SBL 14645
Rushabha Balaji, Kuan-Lin Chen, Bhaskar D. Rao

Spatiotemporal Causal Decoupling Model for Air Quality Forecasting 14650
Jiaming Ma, Guanjun Wang, Sheng Huang, Kuo Yang, Binwu Wang, Pengkun Wang, Yang Wang

Contrastive Pre-Training and Post-Tuning for Heterogeneous Graph Learning 14655
Yulan Hu, Sheng Ouyang, Zhirui Yang, Yong Liu

Unsupervised Motion-Robust Self-Distillation Framework for Remote Physiological Measurement 14660
Anbang Liu, Shanlin Xiao, Wenming Zheng

PromptArtisan: Multi-Instruction Image Editing in Single Pass with Complete Attention Control 14665
Kunal Swami, Raghu Chittersu, Pranav Adlinge, Rajeev Irny, Shashavali Doodekula, Alok Shukla

CPSNet: Comprehensive Enhancement Representation for Polyp Segmentation Task 14670
Jiati Cai, Xiaogang Liu, Hongjie Yang, Yi Ding, Ting Zhong, Zhen Qin

BIGFR: Bridging Individual and Group Fairness in Recommendation Systems 14675
Yaorui Gan, Xuemin Wang, Tieyuan Liu, Liang Chang, Qicang Gen, Yu Zeng

MDNet: Multi-Decoder Network for Abdominal CT Organs Segmentation 14680
Debesh Jha, Nikhil Kumar Tomar, Koushik Biswas, Gorkem Durak, Matthew Antalek, Zheyuan Zhang, Bin Wang, Md Mostafijur Rahman, Hongyi Pan, Alpay Medetalibeyoglu, Vandan Gorade, Abhijit Das, Yury Velichko, Daniela Ladner, Amir Borhani, Ulas Bagci

Pipeline-Centered Neighboring Network for Deep Unfolding Pansharpening 14685
Yan Li, Qiuju Chen, Chuangjie Fang, Ni Xu, Honghui Xu, Jianwei Zheng

Intra- And Inter-Modal Context Interaction Modeling for Conversational Speech Synthesis 14690
Zhenqi Jia, Rui Liu

Towards Expressive Video Dubbing with Multiscale Multimodal Context Interaction 14695
Yuan Zhao, Rui Liu, Gaoxiang Cong

Transformer-Enhanced Iterative Feedback Mechanism for Polyp Segmentation 14700
Nikhil Kumar Tomar, Debesh Jha, Koushik Biswas, Ulas Bagci

Climate Downscaling Using Neural Operator: Spatiotemporal Multimodal Fusion Operator with State-Query Coupled Kernel 14705
Haodi Zhang, Yichi Wang, Yifan Jian, Jiahui Jiang, Zhaohai Bai, Lin Ma

JELLY: Joint Emotion Recognition and Context Reasoning with LLMs for Conversational Speech Synthesis 14710
Jun-Hyeok Cha, Seung-Bin Kim, Hyung-Seok Oh, Seong-Whan Lee

Neural Directed Speech Enhancement with Dual Microphone Array in High Noise Scenario 14715
Wen Wen, Qiang Zhou, Yu Xi, Haoyu Li, Ziqi Gong, Kai Yu

JANE: Joint Angle Networks Assisting 3D Human Pose Estimation 14720
Jinhuan Wang, Yuzhen Zhao, Xujie Song, Wenzhou Chen, Qi Xuan

InvGS: A Novel Real-Time Inverse Rendering Framework Utilizing 3D Gaussian Splatting 14725
Yang Pu, Qingfeng Wu

Explicit Spatial Hint and Implicit Logits Relation: Distilling Heterogeneous Knowledge from Vision Transformer to CNN 14730
Wenkuan Li, Dongyuan Li, Shuo Gao, Xubin Wu, Ming Li, Haifang Li

FLowHigh: Towards Efficient and High-Quality Audio Super-Resolution with Single-Step Flow Matching 14735
Jun-Hak Yun, Seung-Bin Kim, Seong-Whan Lee

FuzzyMIL: Decoupling Pathological Phenotypes Through Deep Fuzzy Clustering for Efficient Whole Slide Image Analysis 14740
Anran Liu, Tong Li, Jing Cai, Srinivasa Sampath Veer Vajrala

Low-Light Detector Based on Feature Filtering and Enhancement 14745
Xinyu Wang, Baoguo Wei, Yuetong Su, Xu Li, Lixin Li

Semantic Data Augmentation for Few-Shot Biomedical Named Entity Recognition 14750
Ying Zhang, Weihua Wang

A Model Stealing Attack Against Multi-Exit Networks 14755
Pan Li, Peizhuo Lv, Kai Chen, Shengzhi Zhang, Yuling Cai, Fan Xiang

On Decentralized Learning with Stochastic Subspace Descent 14760
Shivangi Dubey Sharma, Pranay Sharma, Ketan Rajawat

Towards Context-Aware EEG-Based Emotion Recognition Models: Personality and Emotional Intelligence as Context 14765
Kannadasan K, Nikita Rajesh Verma, Jainendra Shukla

SelfDeblur with Sparsity Enforced Bregman Learning 14770
Yaoyun Zeng, Beier Chen, Hongxia Wang

CAMDet: Condition-Adaptive Multispectral Object Detection Using a Visible-Thermal Translation Model 14775
Junbo Jang, Jiyoon Lee, Joonki Paik

Maximum Mutual Information Estimation Based Graph Attention Network for Knowledge Graph Completion 14780
Wenbin Zhang, Shimei Luo, Zechen Meng, Mankun Zhao, Tianyi Xu, Jian Yu, Jiale Mei, Mei Yu

Enhancing Image Generation Fidelity Via Progressive Prompts 14785
Zhen Xiong, Yuqi Li, Chuanguang Yang, Tiao Tan, Zhihong Zhu, Siyuan Li, Yue Ma

Self-Supervised Uncertainty-Guided Refinement for Robust Joint Optical Flow and Depth Estimation 14790
Rokia Abdein, Wei Li, Chenghao Li, Xiangping Zheng, Rahul Yadav

Accurate 3D Facial Paralysis Analysis Using Multi-View Infrared Structured Light System 14795
Di Wu, Yuping Ye, Jixin Liang, Shiyang Long, Zhan Song

SoCov: Semi-Orthogonal Parametric Pooling of Covariance Matrix for Speaker Recognition 14800
Rongjin Li, Weibin Zhang, Dongpeng Chen, Jintao Kang, Xiaofen Xing

Band Prompting Aided SAR and Multi-Spectral Data Fusion Framework for Local Climate Zone Classification 14805
Haiyan Lan, Shujun Li, Mingjie Xie, Xuanjia Zhao, Hongning Liu, Pengming Feng, Dongli Xu, Guangjun He, Jian Guan

AdapFed: Adaptive Devices Training Strategy for Heterogeneous Federated Learning 14810
Ziyang Chen, Ying Zhou, Fei Song

To Learn Better Character Embeddings in Generative Models for Password Attack 14815
Mingli Zheng, Haibo Cheng, Jiahong Yang, Wenbo Zhang, Ping Wang

DiffSR: Learning Radar Reflectivity Synthesis Via Diffusion Model from Satellite Observations 14820
Xuming He, Zhiwang Zhou, Wenlong Zhang, Xiangyu Zhao, Hao Chen, Shiqi Chen, Lei Bai

Leveraging Pre-Trained Models for Multimodal Class-Incremental Learning Under Adaptive Fusion 14825
Yukun Chen, Zihuan Qiu, Fanman Meng, Hongliang Li, Linfeng Xu, Qingbo Wu

Towards Adversarial Robustness and Backdoor Mitigation in SSL 14830
Aryan Satpathy, Nilaksh Nilaksh, Dhruva Abhijit Rajwade, Somesh Kumar

Federated Cross-Client Collaborative Filtering with Tensor Compressive Learning 14835
Maolan Zhang, Di Xiao, Lvjun Chen, Jindong Xia, Zhuyan Yang

MetaCon: Revitalizing Internet Congestion Control with Meta-Reinforcement Learning..... 14840
He Bai, Hui Li, Jianming Que, Minglong Zhang, Peter Han Joo Chong, Kalupahana Liyanage Kushan Sudheera, Xinyuan Pei

Frequency Agnostic Tissue Characterization in Ultrasound Imaging Using Backscattered Signal Statistics 14845
Abhinav Gadge, Abhishek Kumar, Debdoot Sheet

Adaptive Time-Frequency Attention Network for Sleep Stage Classification Using Respiratory Signals 14850
Zihang Liang, Kejing He

A Clinical Knowledge-Driven Fine-Tuning Strategy for Applying Foundation Model to Fully Automatic Acute Ischemic Stroke Lesion Segmentation on Non-Contrast CT Scans..... 14855
Xianzhen Tan, Zhe Qu, Jie Wang, Hulin Kuang

ARIG-GCN: Anatomical Relationship and Isomorphic Graph Approximation Guided Graph Convolutional Network for Automated ASPECTS Scoring on Non-Contrast CT 14860
Ning Li, Zhe Qu, Jie Wang, Hulin Kuang

MRANet: An Encoder-Decoder Network with Multi-Scale Residual Atrous-Spatial Pyramid Pooling for Seismic Phase Picking..... 14865
Feng Jiang, Hongxi Wei, Yingyue Jing, Lingguo Meng

DLM-VMTL:A Double LayerMapper for Heterogeneous Data Video Multi-Task Prompt Learning..... 14870
Zeyi Bo, Ye Jin, Wuxi Sun

GDFDNet: A Novel Graph-Based Dynamically Fused Dual-Stream Network for Accuracy Prohibited Items Detection 14875
Xiaomeng Li, Hongxia Gao, Yaobin Huang, Zhenming Guan, Litao Li

Improving Cross-Lingual Phonetic Representation of Low-Resource Languages Through Language Similarity Analysis 14880
Minu Kim, Kangwook Jang, Hoirin Kim

Code Drift: Towards Idempotent Neural Audio Codecs..... 14885
Patrick O'Reilly, Prem Seetharaman, Jiaqu Su, Zeyu Jin, Bryan Pardo

Dynamic Routing and Calibration for Few-Shot Object Detection..... 14890
Jiaqi Wu, Jie Lei, Hao Tian, Xiaoqiang Liu, Zunlei Feng, Ronghua Liang

Masked Image Pretraining on Language Assisted Representation 14895
Zejiang Hou, Sun-Yuan Kung

Efficient Estimation of Kernel Matrix Spectral Norm Using Random Features 14900
Yiting Cao, Shayan Shafaei, Luyuan Yang, Chao Lan

Weak-To-Strong Generalization in Speech Recognition 14905
Soheil Khorram, Qian Zhang, Rohit Prabhavalkar, Kartik Audhkhasi, Bhuvana Ramabhadran

Semantics-Guided Dynamic Hypergraph Network for Human Mobility Nowcasting in Disaster 14910
Bowen Zhang, Yunlong Xing, Zinao Su, Jinzhou Cao, Tianhong Zhao, Genan Dai

Zero-Shot Stance Detection with Logically Consistent Data Augmentation 14915
Bowen Zhang, Xu Li, Jun Ma, Xi Zhang, Genan Dai, Jianhua Ye

Fresh-CL: Feature Realignment Through Experts on Hypersphere in Continual Learning 14920
Zhongyi Zhou, Yaxin Peng, Pin Yi, Minjie Zhu, Chaomin Shen

Short-Time Quantum Fourier Transform Processing 14925
Sreeraj Rajindran Nair, Benjamin Southwell, Christopher Ferrie

Zero-Shot Cross-Domain Slot Filling with Retrieval Augmented In-Context Learning 14930
Mengxiao Song, Tingwen Liu, Quangang Li, Duohe Ma, Ming Sun, Ling Tian

Quad-Net: Melspectrogram Vocoder with Convolutional Layers Restricted by the Quadrature Mirror Filter for Perfect Reconstruction 14935
Nam-Seok Song, Joon-Hyuk Chang

HYMAN: Hybrid Memory and Attention Network for Unsupervised Anomaly Detection 14940
Jiahao Li, Yiqiang Chen, Yunbing Xing, Yang Gu, Xiangyuan Lan

E1 TTS: Simple and Fast Non-Autoregressive TTS 14945
Zhijun Liu, Shuai Wang, Pengcheng Zhu, Mengxiao Bi, Haizhou Li

Prototype Matching with Domain Alignment for Open-World Specific Emitter Identification 14950
Wang Xiao, Yalan Ye, Tongjie Pan, Chenyang Li

Investigation of Spatial Self-Supervised Learning and Its Application to Target Speaker Speech Recognition 14955
Yoshiaki Bando, Samuele Cornell, Satoru Fukayama, Shinji Watanabe

Diff4Steer: Steerable Diffusion Prior for Generative Music Retrieval with Semantic Guidance 14960
Xuchan Bao, Judith Yue Li, Zhong Yi Wan, Kun Su, Timo Denk, Joonseok Lee, Dima Kuzmin, Fei Sha

Deep Model Pruning Without Finetuning for Few Category Datasets 14965
Yuchen Huang, Jie Xie, Cheng Wu

TransVDM: Motion-Constrained Video Diffusion Model for Transparent Video Synthesis 14970
Menghao Li, Zhenghao Zhang, Junchao Liao, Long Qin, Weizhi Wang

Data-Efficient Low-Complexity Acoustic Scene Classification Via Distilling and Progressive Pruning 14975
Bing Han, Wen Huang, Zhengyang Chen, Anbai Jiang, Pingyi Fan, Cheng Lu, Zhiqiang Lv, Jia Liu, Wei-Qiang Zhang, Yanmin Qian

Controllable Generative Model for Brain Evolution 14980
Gengshuo Liu, Nikhil N. Chaudhari, Nikos Kanakaris, Chenzhong Yin, Paul Bogdan, Andrei Irimia

DreamHA: Towards High-Quality Human Animation with Image-To-Video Diffusion Models 14985
Longran Shao, Bonan Li, Congying Han, Wenzhao Liu, Tiande Guo, Tianchi Xing, Xinmin Qiu, Zicheng Zhang

Grey Wolf Optimizer Algorithm Based Active Noise Control Without Secondary Path Identification 14990
Zhehua Duan, Tian Zhang, Fei Xu, Chenlin Lu

HiRes: Hierarchical Feature Optimization and Rescorer for Automatic ICD Coding 14995
Zhenpeng Liang, Hongjiao Guan, Wenpeng Lu, Xueping Peng, Bing Xu, Muyun Yang

TriDE-Net: Triple-Densely Extraction Network for Precise Skin Lesion Segmentation 15000
Huan Wan, Taona Deng, Wujian Xu, Xin Wei, Jinshan Zeng

RAFDet: A Novel Camera-Radar Fusion Framework for Robust 3D Object Detection in Autonomous Driving 15005
Xingjian Cao, Ping Wang, Zhitao Zhang, Huizhao Tu, Yong Chen, Zhenbao Liang

Pushing Wi-Fi Towards Fine-Grained Sensing Via Spectrogram Enhancement 15010
Hongbo Jiang, Yiwei Chen, Jingyang Hu, Siyu Chen

Efficient Prototypical Classifier for Class-Incremental Learning 15015
Wei Zhang, Jingyang Qiao, Yuan Xie, Zhizhong Zhang, Xin Tan

Adaptive Prototype Learning for Anomalous Sound Detection with Partially Known Attributes 15020
Anbai Jiang, Xinhu Zheng, Bing Han, Yihong Qiu, Pingyi Fan, Wei-Qiang Zhang, Cheng Lu, Jia Liu

DKD2L: Dual Knowledge Distillation Dynamic Learning for Sketch-Based 3D Shape Retrieval 15025
Yawen Su, Jing Bai, Gan Lin

A Teacher Action Quality Assessment Method Based on Label Constraint Strategy 15030
Ming Fang, Yunpeng Zhou, Jianping Ren, Chunsheng Qin, Shuhua Liu

ConcealGS: Concealing Invisible Copyright Information in 3D Gaussian Splatting 15035
Yifeng Yang, Hengyu Liu, Chenxin Li, Yining Sun, Wuyang Li, Yifan Liu, Yiyang Lin, Yixuan Yuan, Nanyang Ye

RAC-GAN: Iterative Dual-Objective Over and Under Sampling for Imbalanced Datasets 15040
Nitin Gupta, Shashank Mujumdar, Ruhi Sharma Mittal, Hima Patel

Hyperbolic PHATE: Visualizing Continuous Hierarchy of Latent Differentiation Structures 15045
Masahiro Nakano, Hiroki Sakuma, Ryo Nishikimi, Kenji Komiya, Tomoharu Iwata, Kunio Kashino

Improving Robustness of Diffusion-Based Zero-Shot Speech Synthesis Via Stable Formant Generation 15050
Changjin Han, Seokgi Lee, Gyuhyeon Nam, Gyeongsu Chae

Cluster-Perceptive Graph Contrastive Learning for Community Detection 15055
HongBin Li, Fanyu Han, Wei Wang

Graph-Based Signal Sampling with Adaptive Subspace Reconstruction for Spatially-Irregular Sensor Data 15060
Darukeesan Pakiyarajah, Eduardo Pavez, Antonio Ortega

Density-Adaptive Fuzzy Clustering with Isolation Kernel 15065
Paritosh Tiwari, Rankit Kachroo, Punit Rathore

Inter- And Intra-Sentence Cuer-Invariant Representation Learning for Generalizable Cued Speech Recognition 15070
Tianxin Xie, Li Liu

Estimating Multi-Chirp Parameters Using Curvature-Guided Langevin Monte Carlo 15075
Sattwik Basu, Debottam Dutta, Yu-Lin Wei, Romit Roy Choudhury

Large Language Model-Empowered Adversarial Fusion for Typhoon Track Prediction 15080
Lei Luo, Yang Lei, Jiahao Luan, Anudeep Vurity, Sai Sumanth Sriram, Jun Guo

Multi-Object Detection and Tracking Algorithm for Fry Counting Based on DV-YOLO and FryMOT 15085
Hang Yuan, Zhibin Yu, Qiusheng Li, Tianning Fu, Bing Zheng

MemKD: Memory-Discrepancy Knowledge Distillation for Efficient Time Series Classification..... 15090
Nilushika Udayangani, Kishor Nandakishor, Marimuththu Palaniswami

Security-Enhanced Data Transmission Scheme for IoT-Based Healthcare in Remote Areas..... 15095
Zhenbin Guo, Yuchuan Luo, Shaojing Fu, Ming Xu

Extracting Sparse Specialist Models from Generalist Models 15100
Tao Yu, Xu Zhao, Yongqi An, Guibo Zhu, Ming Tang, Jinqiao Wang

High-Resolution Gait Micro-Doppler Synthesis from Videos Over Diverse Trajectories..... 15105
Shubo Yang, Soheil Hor, Jae-Ho Choi, Amin Arbabian

SemiGPS: GraphGPS-Based Semi-Supervised Graph Learning for Sector-Specific GDP Mapping.....15110
Jinzhou Cao, Xiangxu Wang, Jiashi Chen, Bowen Zhang, Yahan Ma, Tianhong Zhao

Semantic Prior-Guided Scalable Image Coding 15115
Wuzhen Shi, Wennan Yin, Fei Tao, Yang Wen

Class Semantic Prompts Enhanced Prototypical Fusion Method for Few-Shot Named Entity Recognition 15120
Mei Yu, Yuang Tao, Mankun Zhao, Tianyi Xu, Zechen Meng, Wenbin Zhang, Jian Yu

Exact Solutions of the Inner Optimization Problem of Adversarial Robustness 15125
Deepak Maurya, Adarsh Barik, Jean Honorio

A Parametric Non-Negative Coupled Canonical Polyadic Decomposition Algorithm for Hyperspectral Super-Resolution..... 15130
Xi-Yuan Liu, Xiao-Feng Gong, Lei Wang, Wei Feng, Qiu-Hua Lin

Target Localization with a Coprime Multistatic MIMO Radar Via Coupled Canonical Polyadic Decomposition Based on Joint EVD 15135
Guo-Zhao Liao, Xiao-Feng Gong, Wei Liu, Hing Cheung So

Facial Expression Recognition with DToF Sensing 15140
Chengxiao Li, Xie Zhang, Chenshu Wu

AIDC: Benchmark for Analytical Learning in Incremental Disease Classification 15145
Rongchang Zhao, Jianyu Qi, Rui Li, Zhijie Zheng, Jian Zhang, Jiaxu Li

Semantic Attention and LLM-Based Layout Guidance for Text-To-Image Generation..... 15150
Yuxiang Song, Zhaoguang Long, Man Lan, Changzhi Sun, Aimin Zhou, Yuefeng Chen, Hao Yuan, Fei Cao

Hyper-Adapter for Parameter-Efficient Multilingual ASR Adaptation 15155
Zejiang Hou, Daniel Garcia-Romero, Kyu J. Han

A 3D Attenuation Coefficient Based Degradation Estimation for Real Non-Homogeneous Dehazing 15160
Hanyu Jiang, Zemin Ren, Tao Tan

Uniform Distribution Based Learnable Quantization Via Self-Knowledge Distillation..... 15165
Peng Hu, Biao Leng

A Block Term Decomposition Model Based Algorithm for Tensor Completion of Multidimensional Harmonic Signals 15170
Lei Wang, Xiao-Feng Gong, Xi-Yuan Liu, Wei Feng, Qiu-Hua Lin

Sound Source Distance Estimation Utilizing Physics-Informed Prior for Sound Event Localization and Detection 15175
Nao Sato, Masahiro Yasuda, Shoichiro Saito, Noboru Harada

Brain MRI Segmentation with Language-Driven Detection and Context-Aware Descriptions 15180
Qiang Fu, Xinyuan Xia, Yi Hong

Variational Perturbation Personalized Federated Learning Via Prior-Posterior Distance 15185
Hefeng Zhou, Yuanbin Wang, Jun Wang, Jiong Lou, Wugedele Bao, Chentao Wu, Jie Li

Low-Rank Transformer Adaptation for Arbitrary Style Transfer 15190
Wenjie Xu, Meichen Liu, Bihan Wen

Can RAG-Driven Enhancements Amplify Audio LLMs for Low-Resource Languages? 15195
Bikash Dutta, Rishabh Ranjan, Akshat Jain, Richa Singh, Mayank Vatsa

Cross-Layer Cache Aggregation for Token Reduction in Ultra-Fine-Grained Image Recognition 15200
Edwin Arkel Rios, Jansen Christopher Yuanda, Vincent Leon Ghanz, Cheng-Wei Yu, Bo-Cheng Lai, Min-Chun Hu

Adaptive Password Guessing Framework Using Various Datasets 15205
Wenbo Zhang, Haibo Cheng, Mingli Zheng, Jiahong Yang, Ping Wang

DRSFANet: Dual-Path CNN with Residual and Frequency Attention for Image Denoising 15210
Manish Kumar, Suman Kumar Maji, Anirban Saha

Evaluating the Impact of Discriminative and Generative E2E Speech Enhancement Models on Syllable Stress Preservation 15215
Rangavajjala Sankara Bharadwaj, Jhansi Mallela, Sai Harshitha Aluru, Chiranjeevi Yarra

PatchST: A Patch Spatial-Temporal Network for Large-Scale Traffic Forecasting 15220
Jinrun Li, Gaowei Zhang, Wei Wang, Yi Wang

Generative Adversarial Network with Structured Semantic Prompts Constrainting Clip for Text-To-Image 15225
Shuheng Ge, Li Zhang, Haoyu Xing, Xiangqian Wu

Unified Audio Event Detection 15230
Yidi Jiang, Ruijie Tao, Wen Huang, Qian Chen, Wen Wang

Prelude Echoes Finale: Video Domain Adaptation with Fine-Grained Temporal Consistency 15235
Shuxian Wang, Mengmeng Jing, Xianlong Tian, Yuguo Hu, Lin Zuo

LKSNeXt: An Efficient Medical Image Segmentation Network with Large Kernels and Lightweight Structure 15240
Qingxue Zhao, Di Wu, Ge Tang, Jun Tian

Post-Net2.0: An Adaptive Weighted Loss Function Driven by Linguistic Constraint for Automatic Syllable Stress Detection 15245
Sai Harshitha Aluru, Jhansi Mallela, Chiranjeevi Yarra

InsectMamba: State Space Model with Adaptive Composite Features for Insect Recognition 15250
Qianning Wang, Chenglin Wang, Zhixin Lai, Yucheng Zhou

Situational Awareness Based Resource Allocation for Multi-Target Tracking in Distributed Radar Network 15255
Mushen Lin, Fenggang Yan, Lingda Ren, Xiangtian Meng, Maria Greco, Fulvio Gini, Ming Jin

MambaCPU: Enhanced Correlation Mining with State Space Models for CPU Performance Prediction 15260
Xiaoman Liu

An Improved Planar Approximation Localization Method in Distributed Airborne Radars 15265
Jing Li, Jinshan Ding, Liwu Wen, Zehua Yu, Demin Huang

See in Detail: Enhancing Sparse-View 3D Gaussian Splatting with Local Depth and Semantic Regularization 15270
Zongqi He, Zhe Xiao, Kin-Chung Chan, Yushen Zuo, Jun Xiao, Kin-Man Lam

Spatial Annotation-Free Training for Sound Event Localization and Detection 15275
Masahiro Yasuda, Shoichiro Saito, Nao Sato, Noboru Harada

DeepMatch: Navigating the Complexities of Underwater Textures for Enhanced Keypoint Matching 15280
Sainan Zhang, Tong Liu, Zhe Wang, Zhibin Yu, Bing Zheng

Mutual-View Contrastive Generative Framework for Attribute-Missing Graph Clustering 15285
Shijun Li, Changjian Wang, Kele Xu, Xiaojin Li, Gaojin He, Xu Liu

FCoDT-Net: A Novel Framework for High-Precision Medical Image Segmentation Using Contextual Distillation Transformer 15290
Qin YuTao, Yang SiZhe, Hu Bang, Ren Wei

Improved Bounds for Online Convex Optimization 15295
Tanvi S Nayak, B. N. Bharath

Adapting Large Language Model for Spatio-Temporal Understanding in Next Point-Of-Interest Prediction 15300
Qiuhan Han, Atsushi Yoshikawa, Masayuki Yamamura

Multi-Range Adaptive Perception Transformer for Iterative Homography Estimation 15305
Tianming Li, Qing Zhu, Zhen Zhou, Jianqiao Luo, Yaonan Wang

Integrating Spectro-Temporal Cross Aggregation and Multi-Scale Dynamic Learning for Audio Deepfake Detection 15310
Yunqi Hao, Minqiang Xu, Yihao Chen, Yanyan Liu, Liang He, Lei Fang, Lin Liu

Rethinking Adversarial Attacks in Reinforcement Learning from Policy Distribution Perspective 15315
Tianyang Duan, Zongyuan Zhang, Zheng Lin, Yue Gao, Ling Xiong, Yong Cui, Hongbin Liang, Xianhao Chen, Heming Cui, Dong Huang

Multi-Level Conflict-Aware Network for Multi-Modal Sentiment Analysis 15320
Yubo Gao, Haotian Wu, Lei Zhang

LPBS: A RL-PPO Driven K8S Batch Processing Task Scheduler 15325
Shengchao Yuan, Xiaoxin Bai

MMA-Net: Multi-Modal Attention Network for 2-D Object Detection in Autonomous Driving 15330
Abhilash Gaur, Shubh Goel, Kanishk Goel, Seshan Srirangarajan, Po-Hsuan Tseng, Kai-Ten Feng

Negative Learning and Dual Contrastive for Unsupervised Visible-Infrared Person Re-Identification 15335
Jiajia Xu, Xuemiao Xu, Weiwei Cai

FlowSE: Flow Matching-Based Speech Enhancement 15340
Seonggyu Lee, Sein Cheong, Sangwook Han, Jong Won Shin

Autoregressive Language Model with Historical Context Re-Encoding 15345
Yimeng Zhuang

Multi-Modal Streaming ASR in Cross-Talk Scenario for Smart Glasses 15350
Ya Jiang, Hongbo Lan, Qing Wang, Shutong Niu

SECC-Stega: Generative Linguistic Steganographic Framework Based on Error Correcting Codes 15355
Yuzhe Guo, Zhongliang Yang, Zhuang Wang, Zhili Zhou, Linna Zhou

STGE-Former: Spatial-Temporal Graph-Enhanced Transformer for EEG-Based Major Depressive Disorder Detection 15360
Yu Chen, Chunfeng Yang

Efficient Data-Dependent Random Projection for Least Square Regressions 15365
Jacob Sturges, Luyuan Yang, Shayan Shafaei, Chao Lan

Revolutionizing Disease Diagnosis with Simultaneous Functional PET/MR and Deeply Integrated Brain Metabolic, Hemodynamic, and Perfusion Networks 15370
Luoyu Wang, Yitian Tao, Qing Yang, Yan Liang, Siwei Liu, Hongcheng Shi, Dinggang Shen, Han Zhang

High-Efficiency Modulation Classification with Temporal-Frequency Analysis Based on Multi-Channel Filter Bank 15375
Yifan Dai, Xin Gao, Ke Jing, Bin Tian

Predicting Biomechanical Risk Factors for Division - I Women's Basketball Athletes 15380
Aayushi Shah, Vanaja Agarwal, Dhairya Shah, Harman Jani, Srishti Sharma, Tolga Kaya, Christopher Taber, Mehul S. Raval

Redesigning Graph Filter-Based GNNs to Relax the Homophily Assumption 15385
Samuel Rey, Madeline Navarro, Victor M. Tenorio, Santiago Segarra, Antonio G. Marques

Aesthetic Perception Prompting for Interpretable Image Aesthetics Assessment with MLLMs 15390
Lanjun Wang, Zheyu Qiao, Ruidong Chen, Jingqiu Li, Wenjie Wang, Xiaoqiong Wang, Wei Rao, Shuai Chen, An-An Liu

KIKE: Linguistic Steganalysis Based on Knowledge Infusion and Knowledge Encoding 15395
Zhuang Wang, Xuekai Chen, Zhongliang Yang, Linna Zhou

TrackFusion: Enhancing Multi-Object Tracking with Temporal Trajectory Modeling and Frame-Integrated Detection 15400
Xinhao Zhang, Shuai Liu, Bingyang Wang, Jiaojiao Dai, Jinqing Qi, Huchuan Lu, You He

DAEF-VS: An Efficient Universal VoIP Steganalysis Framework Based on Domain-Aware Knowledge 15405
ZhengYang Fang, Pengcheng Zhou, Zhongliang Yang, Zhili Zhou, Linna Zhou

CLHi-MTS: A Contrastive Learning-Based Hierarchical Framework for Masked Medical Time-Series Modeling 15410
Ziyang Cheng, Shurong Sheng, Xiongfei Wang, Yi Sun, Kuntao Xiao, Wanli Yang, Pengfei Teng, Guoming Luan, Jiahong Gao

OCLNet: Obfuscation Feature Contrastive Learning Network for Weakly Supervised Semantic Segmentation on Ultrasound Images 15415
Jie Gao, Xianzhi Zhang, Zijian Zhang, Xuewei Li, Mei Yu, Ruiguo Yu, Zhiqiang Liu

ECSNN: Spiking Neural Networks for Efficient Exposure Correction in Endoscopy Imaging 15420
Jun Zhang, Zhuoran Zheng, Jingang Zhang, Wenqi Ren

Point Cloud Resampling with Learnable Heat Diffusion 15425
Wenqiang Xu, Wenrui Dai, Duoduo Xue, Ziyang Zheng, Chenglin Li, Junni Zou, Hongkai Xiong

PTQ4ADM: Post-Training Quantization for Efficient Text Conditional Audio Diffusion Models 15430
Jayneel Vora, Aditya Krishnan, Nader Bouacida, Prabhu R V Shankar, Prasant Mohapatra

Wavelet Scattering Network Features for Intensity Category Classification and Prediction of SPL from Speech 15435
Manila Kodali, Sudarsana Reddy Kadiri, Shrikanth Narayanan, Paavo Alku

Multi-Source Spatial Knowledge Understanding for Immersive Visual Text-To-Speech 15440
Shuwei He, Rui Liu

FedDiffRec: A Module-Wise Training Approach for Diffusion-Based Recommendation in Federated Learning 15445
Guohui Li, Lu Zhang, Qian Rong, Xuanang Ding, Ling Yuan

Efficient and Effective Model Extraction 15450
Hongyu Zhu, Wentao Hu, Sichu Liang, Fangqi Li, Wenwen Wang, Shilin Wang

Two-Stream Spiking Neural Network for Event-Based Action Recognition 15455
Shuang Lian, Qianhui Liu, Ziling Wang, Jia Su, Zhibin Zuo, Yi Zhang, Rui Yan, Huajin Tang

Exploring Spectral Signatures of Chinese Liquor Using Machine Learning and SHapley Additive exPlanations 15460
Danlei Chen, Yun Wang, Linruize Tang, Zhengqiao Zhao, Jie Chen, Jingdong Chen

GMCL: Graph-Enhanced Multimodal Contrastive Learning for Rumor Detection 15465
Kun Lu, Hongli Zhang, Tianze Sun, Yuchen Yang, Chao Meng, Gongzhu Yin, Binxing Fang

WTU-EVAL: A Whether-Or-Not Tool Usage Evaluation Benchmark for Large Language Models 15470
Jian Liu, Kangyun Ning, Yisong Su, Wenjuan Han, Jinan Xu, Yuanzhe Zhang

Channel and Space-Based Joint Rate Allocation Algorithm 15475
Dayong Wang, Chao Yuan, Yu Sun, Xin Lu, Hui Guo, Frederic Dufaux, Ce Zhu

3D Shape Classification by Registration: Neural-Network-Free and Training-Free 15480
Chang Gou, Yuanqu Mou, Wenjie Li, Neetesh Purohit, Suneel Yadav, Haiyang Bai, Xu Zhang, Lijun Chen

Toward Robust Early Detection of Alzheimer's Disease Via an Integrated Multimodal Learning Approach 15485
Yifei Chen, Shenghao Zhu, Zhaojie Fang, Chang Liu, Binfeng Zou, Linwei Qiu, Yuhe Wang, Shuo Chang, Fan Jia, Feiwei Qin, Jin Fan, Yong Peng, Changmiao Wang

Egocentric Speaker Diarization with Vision-Guided Clustering and Adaptive Speech Re-Detection..... 15490
He Huang, Haoyuan Yu, Daibo Liu, Haowen Chen, Minjie Cai

FSENet: Frequency Separation Enhancement Network for Super-Resolution..... 15495
Liangdong Li, Zhihuai Xie

HyperSDT: HyperNetwork Slide Decision Tree for Interpretable Tabular Learning..... 15500
Nan Hu, Xueqiong Li, Jun-Jie Huang, Zhenhua Liang, Shaowu Yang, Ji Wang

GRACED: A Plug-And-Play Solution for Certifiable Graph Classification..... 15505
Xiaoyu Liang, Haohua Du, He Lu, Fei Shang

Partial Inference in Structured Prediction..... 15510
Chuyang Ke, Deepak Maurya, Jean Honorio

TIDE-Net: A Physics-Based Graph Model for Predicting Tropical Cyclone Impacts on Estuarine Systems..... 15515
Gaowei Zhang, Wei Wang, Yi Wang

Lightweight Image Quality Prediction Guided by Perceptual Ranking Feedback..... 15520
Aymen Sekhri, Mohamed-Chaker Larabi, Seyed Ali Amirshahi

HyperMST: Multi-Scale Spatio-Temporal Hypercorrelation Network for POI Recommendation..... 15525
Zeyun Zhao, Changjian Wang, Kele Xu, Zhen Huang, Gaojin He, Xu Liu

SSVEP-BiMA: Bifocal Masking Attention Leveraging Native and Symmetric-Antisymmetric Components for Robust SSVEP Decoding..... 15530
Yuxin Liu, Zhenxi Song, Guoyang Xu, Zirui Wang, Feng Wan, Yong Hu, Min Zhang, Zhiguo Zhang

M3Rec: Selective State Space Models with Mixture-Of-Modality Experts for Multi-Modal Sequential Recommendation..... 15535
Xu Guo, Tong Zhang, Yufei Xue, Chenxu Wang, Fuyun Wang, Zhen Cui

EEG-ReMinD: Enhancing Neurodegenerative EEG Decoding Through Self-Supervised State Reconstruction-Primed Riemannian Dynamics..... 15540
Zirui Wang, Zhenxi Song, Yi Guo, Yuxin Liu, Guoyang Xu, Min Zhang, Zhiguo Zhang

StrucFormer: Structural Prior Guided Transformer for Mobile Crowdsensing Data Inference..... 15545
Xu Kang, Shouceng Tian, Feifei Kou, Lei Shi, Jiadong Ren

FedCAda: Adaptive Client-Side Optimization for Accelerated and Stable Federated Learning..... 15550
Liuzhi Zhou, Yu He, Kun Zhai, Xiang Liu, Sen Liu, Xingjun Ma, Guangnan Ye, Hongfeng Chai

CSD: Weather Forecasting with Graph Neural Network Based on Cross-Scale Diffusivity..... 15555
Jinrun Li, Gaowei Zhang, Wei Wang, Yi Wang

Factorized-VITS: Decoupling Prosody and Text in End-To-End Speech Synthesis Without External Or Secondary Aligner..... 15560
Yining Liu, Alexander Waibel

LHQ-SVC: Lightweight and High Quality Singing Voice Conversion Modeling..... 15565
Yubo Huang, Xin Lai, Muyang Ye, Anran Zhu, Zixi Wang, Jingzehua Xu, Shuai Zhang, Zhiyuan Zhou, Weijie Niu

Efficient Spatial Audio Rendering Via Differentiable FIR to IIR Estimation..... 15570
Armin Gerami, Bowen Zhi, Dmitry N. Zotkin, Ramani Duraiswami

Meta-Analogy Learning Based on Dynamic Graph Neural Networks for Inductive Knowledge Graph Link Prediction 15575
Jingyu Wang, Zhijuan Du, Tao Sun

Self-Geometry-Guided Direct Pose Regression Based on Dual Perspective Fusion for 2D-3D Cross Dimensional Spinal Surgery Navigation 15580
Jing Ling, Zhengyang Wu, Changqing Li, Weisheng Li, Chao Zhang, Yucheng Shu

VisQ2SQL: Towards SQL-Driven Data Visualization Via LLMs-Grounded Preference Learning..... 15585
Shengze Shi, Tao Ren, Jun Hu

Stable Audio Open..... 15590
Zach Evans, Julian D. Parker, CJ Carr, Zack Zukowski, Josiah Taylor, Jordi Pons

Integrating Multi-Scale Compression Attention with Edge Detection for Ultrasound Tumor Segmentation 15595
Ruoshi Li, Hao Qi, Xing Chen, Yinran Chen

Triple Path Enhanced Neural Architecture Search for Multimodal Fake News Detection..... 15600
Bo Xu, Qiujie Xie, Jiahui Zhou, Linlin Zong

TMANet: Triple Multi-Scale Attention Based Network with Boundary Association Loss for Superpixel Segmentation..... 15605
Ziyi Zhang, Shijie Lian, Hua Li

Heterogeneous Packet Translation for Cross-Technology Communication 15610
Hao He, Qihuan Wu, Shuo Yuan, Ziyin Gu, Zhipeng Yu, Qingmeng Zhu

Evidential Deep Learning with Reweighted Margin Adjustment for Uncertainty-Driven Cervical OCT Image Diagnosis 15615
Hanfeng Zhu, Yi Qu, Yutao Ma, Yuchen Pei

PAFedMIS: Personalized Asynchronous Federated Learning for Medical Image Segmentation..... 15620
Yi Li, Yue Hua, Xin Zheng, Yanqing Guo, Bo Wang

PsSR: Hybrid Path Selection Mechanism for Efficient Image Super-Resolution..... 15625
Cheng Ding, ZhongQiu Zhao, Yang Zhao

DCASI: A Sequence-Based Attack Investigation Method Using DTW Contrastive Learning..... 15630
Yun Li, Wei Qiao, Yan Zhu, Yunxiang Wang, Bo Jiang, Zhigang Lu

Fusing Multimodality of Large Language Models and Satellite Imagery Via Simplicial Contrastive Learning for Latent Urban Feature Identification and Environmental Application..... 15635
Yuzhou Chen, Jiue-An Yang, Hugo Kyo Lee, Calvin Tribby, Tarik Benmarhnia, Marta Jankowska, Yulia R. Gel

Towards Dynamic Neural Communication and Speech Neuroprosthesis Based on Viseme Decoding 15640
Ji-Ha Park, Seo-Hyun Lee, Soowon Kim, Seong-Whan Lee

CoMT: Chain-Of-Medical-Thought Reduces Hallucination in Medical Report Generation..... 15645
Yue Jiang, Jiawei Chen, Dingkang Yang, Mingcheng Li, Shunli Wang, Tong Wu, Ke Li, Lihua Zhang

Targeted Password Guessing Using Neural Language Models 15650
Jiahong Yang, Wenting Li, Haibo Cheng, Ping Wang

DA-LIF: Dual Adaptive Leaky Integrate-And-Fire Model for Deep Spiking Neural Networks 15655
Tianqing Zhang, Kairong Yu, Jian Zhang, Hongwei Wang

Towards Precision Characterization of Communication Disorders Using Models of Perceived Pragmatic Similarity 15660
Nigel G. Ward, Andres Segura, Georgina Bugarini, Heike Lehnert-LeHouillier, Dancheng Liu, Jinjun Xiong, Olac Fuentes

Knowledge Transfer Across Modalities for Weakly Supervised Point Cloud Semantic Segmentation 15665
Zihan Wang, Yunhang Shen, Mengtian Li, Ke Li, Xing Sun, Shaohui Lin, Lizhuang Ma

Enhancing 6D Pose Estimation with Cross-Modal Fusion Network and Density-Peak Keypoint Localization 15670
Liming Zhang, Qing Li, Zhenhong Chen, Chuan Yan, Xiaojiang Peng

Noise-Mitigated Variational Quantum Eigensolver with Pre-Training and Zero-Noise Extrapolation 15675
Wanqi Sun, Jungang Xu, Chenghua Duan

Multi-Layer Network Disintegration Via Deep Reinforcement Learning 15680
Zhenhua Liang, Xueqiong Li, Jun-Jie Huang, Nan Hu, Shaowu Yang, Hengzhu Liu

MQAD: A Large-Scale Question Answering Dataset for Training Music Large Language Models 15685
Zhihao Ouyang, Ju-Chiang Wang, Daiyu Zhang, Bin Chen, Shangjie Li, Quan Lin

Hierarchical Loss for Bi-Level Classification of Speech into Language and Dialects 15690
Ananya Angra, Muralikrishna H, Dileep A. D., Veena Thenkanidiyoor

Easing Optimization Paths: A Circuit Perspective 15695
Ambroise Odonnat, Wassim Bouaziz, Vivien Cabannes

Medical Image Segmentation with Auxiliary Points Prediction of Lesion Location and Boundary 15700
Yu Fan, Zhihui Lai, Heng Kong, Tianying Feng

Bilingual Dual-Head Deep Model for Parkinson's Disease Detection from Speech 15705
Moreno La Quatra, Juan Rafael Orozco-Arroyave, Marco Sabato Siniscalchi

SELMA: A Speech-Enabled Language Model for Virtual Assistant Interactions 15710
Dominik Wagner, Alexander Churchill, Siddharth Sigtia, Erik Marchi

Exploring the Robustness of In-Context Learning with Noisy Labels 15715
Chen Cheng, Xinzhi Yu, Haodong Wen, Jingsong Sun, Guanzhang Yue, Yihao Zhang, Zeming Wei

FABLE: A Bundle Method for Federated Learning in Wireless Systems 15720
Daniel Cederberg, Erik G. Larsson, Mikael Johansson

Adaptive Multi-Scale Local Correction for Semi-Supervised 3D Medical Image Segmentation 15725
Xinqiang Wang, Wenhuan Lu, Ke Zheng, Junhai Xu

Disentanglement Analysis in Deep Latent Variable Models Matching Aggregate Posterior Distributions 15730
Surojit Saha, Sarang Joshi, Ross Whitaker

Segmentation-Guided Sparse Transformer for Under-Display Camera Image Restoration 15735
Jingyun Xue, Tao Wang, Pengwen Dai, Kaihao Zhang

Predicting Compact Phrasal Rewrites with Large Language Models for ASR Post Editing 15740
Hao Zhang, Felix Stahlberg, Shankar Kumar

Learning Simultaneous Facial Canonical Correlation Representation for Face Hallucination..... 15745
Yun-Hao Yuan, Jin Li, Jipeng Qiang, Yi Zhu, Xiaobo Shen, Yun Li

Learning with Partial Labels from Conflict-Free and Semi-Supervised Perspective* 15750
Chen-Chen Zong, Sheng-Jun Huang

Improvised Performance Following in Real Time for Automatic Accompaniment..... 15755
Junyan Jiang, Akira Maezawa, Gus Xia

UniPlane: Unified Plane Detection and Reconstruction from Posed Monocular Videos 15760
Yuzhong Huang, Fred Morstatter

OrientDream: Streamlining Text-To-3D Generation with Explicit Orientation Control 15765
Yuzhong Huang, Fred Morstatter

Dynamic Object Queries for Transformer-Based Incremental Object Detection 15770
Jichuan Zhang, Wei Li, Shuang Cheng, Yali Li, Shengjin Wang

Sequential DOA Trajectory Estimation Using Deep Complex Network and Residual Signals..... 15775
Shreyas Jaiswal, Peter Gerstoft, Santosh Nannuru

Modular Gaussian Splatting: Instance Decomposable Learning and Adaptive Rendering of 3D Scenes Via Mixture of Experts 15780
Jiansong Sha, Haoyu Zhang, Qiangjuan Huang, Guang Kou

Self-Enhanced Reasoning Training: Activating Latent Reasoning in Small Models for Enhanced Reasoning Distillation 15785
Yong Zhang, Bingyuan Zhang, Zhitao Li, Ming Li, Ning Cheng, Minchuan Chen, Tao Wei, Jun Ma, Shaojun Wang, Jing Xiao

Local Adaptive Time-Frequency Bidirectional Synchrosqueezing Transform..... 15790
Cuiwentong Xu, Yuhe Liao

Exploring the Distribution of Cell Subpopulations in Pancreatic Ductal Adenocarcinoma Slides by Joint Spatial Transcriptomics and Pathology Data 15795
Yaqi Deng, Wenjie Cai, Bentao Song, Bin Yang, Lingming Kong, Qingfeng Wang, Jun Huang

Improving Irregular Text Recognition with Adaptive Feature Compression..... 15800
Yin Liu, Zhineng Chen

Boosting Large Language Model for Speech Synthesis: An Empirical Study 15805
Hongkun Hao, Long Zhou, Shujie Liu, Jinyu Li, Shujie Hu, Rui Wang, Furu Wei

A Unified Joint Contrastive Triplet Loss with Temporal and Frequency Signal Fusion for Diagnosing Heart Murmurs..... 15810
Ayushi Pal, Arka Roy, Udit Satija

Patch Attention Excitation Based Vision Transformer for Small-Sized Datasets..... 15815
Akash Verma, Shiv Ram Dubey, Satish Kumar Singh

Adopting Whisper for Confidence Estimation 15820
Vaibhav Aggarwal, Shabari S Nair, Yash Verma, Yash Jogi

Semi-Supervised Video Anomaly Detection with Compact Deformable 3D Convolution 15825
Shibo Gao, Peipei Yang, Linlin Huang

Fully Connected Tensor Network Based Brain Structural Feature Extraction for Early Alzheimer's Disease Detection 15830
Fei He, Xinyue Li, Ce Zhu, Fan Zhang, Yipeng Liu

Directional Source Separation for Robust Speech Recognition on Smart Glasses 15835
Tiantian Feng, Ju Lin, Yiteng Huang, Weipeng He, Kaustubh Kalgaonkar, Niko Moritz, Li Wan, Xin Lei, Ming Sun, Frank Seide

Positional Differential Encoding for Distributed Learning 15840
Leah Woldemariam, Anna Scaglione

Self-Structure Enhance Network for Digital Molar Wax-Up Design 15845
Yuze Shi, Yiqing Wang, Haisheng Li, Li Chen

Domain-Aware Knowledge Debiasing for Generalizable Video Understanding in CLIP 15850
Qingmeng Zhu, Qihuan Wu, Zhipeng Yu, Yi Li, Ziyin Gu, Hao He

SLAM-AAC: Enhancing Audio Captioning with Paraphrasing Augmentation and CLAP-Refine Through LLMs 15855
Wenxi Chen, Ziyang Ma, Xiquan Li, Xuenan Xu, Yuzhe Liang, Zhisheng Zheng, Kai Yu, Xie Chen

SigmoidGS: To Guide Depth More Effectively 15860
Hongai Chow, Licheng Shen, Lingyun Wang, Tong Zhang, Mengqiu Wang, Yuxing Han

Fast DCT+: A Family of Fast Transforms Based on Rank-One Updates of the Path Graph 15865
Samuel Fernández Menduiña, Eduardo Pavez, Antonio Ortega

Towards Bringing Parity in Pretraining Datasets for Low-Resource Indian Languages 15870
Kaushal Santosh Bhogale, Deovrat Mehendale, Tahir Javed, Devbrat Anuragi, Sakshi Joshi, Sai Sundaresan, Aparna Ananthanarayanan, Sharmistha Dey, Sathish Kumar Reddy G, Anusha Srinivasan, Abhigyan Raman, Pratyush Kumar, Mitesh M. Khapra

Hedging is Not All You Need: A Simple Baseline for Online Learning Under Haphazard Inputs 15875
Himanshu Buckchash, Momojit Biswas, Rohit Agarwal, Dilip K. Prasad

Distributed-Robust Source Localization in Wireless Acoustic Sensor Networks 15880
Xu Wang, De Hu, Qintuya Si

Enhancing Out-Of-Distribution Detection Through Dynamic Activation Function 15885
Yingrui Ji, Yao Zhu, Zhigang Li, Jiansheng Chen, Yunlong Kong, Jingbo Chen

M3-CVC: Controllable Video Compression with Multimodal Generative Models 15890
Rui Wan, Qi Zheng, Yibo Fan

ExVC: Leveraging Mixture of Experts Models for Efficient Zero-Shot Voice Conversion 15895
Obed Irihose, Le Zhang

Combining Spatio-Temporal Networks and Graph Attention Architectures for EEG-Based Workload Classification 15900
Nikhil Panwar, Vishal Pandey, Rownak Tiwari, Partha Pratim Roy

A Study of Multi-Scale Feature Learning from Pre-Trained Models on Speaker Verification 15905
Shengyu Peng, Wu Guo, Jie Zhang, Zuoliang Li, Yu Guan, Bin Gu, Yang Ai

Enhanced Control for Diffusion Bridge in Image Restoration 15910
Conghan Yue, Zhengwei Peng, Junlong Ma, Dongyu Zhang

ZCS-CDiff: A Zero-Shot Code-Switching TTS System with Conformer-Based Diffusion Model 15915
Ke Chen, Zhihua Huang, Liang He, Yonghong Yan

Mixture of Experts Fusion for Fake Audio Detection Using Frozen Wav2vec 2.0 15920
Zhiyong Wang, Ruibo Fu, Zhengqi Wen, Jianhua Tao, Xiaopeng Wang, Yuankun Xie, Xin Qi, Shuchen Shi, Yi Lu, Yukun Liu, Chenxing Li, Xuefei Liu, Guanjun Li

Pathological Prior-Guided Multiple Instance Learning for Mitigating Catastrophic Forgetting in Breast Cancer Whole Slide Image Classification 15925
Weixi Zheng, Aoling Huang, Jingping Yuan, Haoyu Zhao, Zhou Zhao, Yongchao Xu, Thierry Géraud

High-Fidelity Stereoscopic Image Rain Removal with Texture Integrity and Disparity Consistency..... 15930
Yanyan Wei, Zhao Zhang, Zhong-Qiu Zhao, Yang Zhao, Richang Hong, Yi Yang, Meng Wang

Person-In-Bed Detection Using Frequency Domain Features and GLR-Based CuSum 15935
Dhinesh Chandran G, Srikrishna Bhashyam, Srinivas Reddy Kota

IPP-Net: A Generalizable Deep Neural Network Model for Indoor Pathloss Radio Map Prediction..... 15937
Bin Feng, Meng Zheng, Wei Liang, Lei Zhang

Probabilistic Person-In-Bed Detection Using Accelerometer Signals 15939
Kaite Shi, Jiayi Gao, Yi Zhou, Xuliang Yu, Liang Zhao

HamaraAwaz: Advancing Low-Latency Streaming TTS for Multilingual Speech in Indian Languages..... 15941
Neil Shah, Parth Khadse, Shirish Karande, Sunil Kumar Kopparapu

Spectral-Temporal Fusion Representation for Person-In-Bed Detection..... 15943
Xuefeng Yang, Shiheng Zhang, Jian Guan, Feiyang Xiao, Wei Lu, Qiaoxi Zhu

GoLoColor: Towards Global-Local Semantic Aware Image Colorization 15945
Tianai Yue, Xiangcheng Du, Jing Liu, Zhongli Fang

Build LLM-Based Zero-Shot Streaming TTS System with Cosyvoice..... 15950
Xiang Lyu, Yuxuan Wang, Tianyu Zhao, Hao Wang, Huadai Liu, Zhihao Du

Mitigating Category Imbalance: Fosafer System for the Multimodal Emotion and Intent Joint Understanding Challenge 15952
Honghong Wang, Yankai Wang, Dejun Zhang, Jing Deng, Rong Zheng

Energy-Based Model Guided Self-Supervised Learning for Speaker Verification..... 15954
Yaqian Hao, Chenguang Hu, Chong Bian, Junlan Feng, Yingying Gao, Shilei Zhang

Enhancing Convolutional Models for Indoor Radio Mapping Via Ray Marching 15959
Mengfan Wu, Marco Skocaj, Mate Boban

HLTCOE Submission to the VoicePrivacy Attacker Challenge..... 15961
Henry Li Xinyuan, Ashi Garg, Zexin Cai, Kevin Duh, Leibny Paola García-Perera, Sanjeev Khudanpur, Nicholas Andrews, Matthew Wiesner

Generalizable Indoor Path Loss Prediction 15963
Cheick Tidiani Cisse, Oumaya Baala, Valéry Guillet, François Spies, Alexandre Caminada

SpecWav-Attack: Leveraging Spectrogram Resizing and Wav2Vec 2.0 for Attacking Anonymized Speech 15965
Yuqi Li, Yuanzhong Zheng, Zhongtian Guo, Yaoxuan Wang, Jianjun Yin, Haojun Fei

Text-Guided Multimodal Fusion for the Multimodal Emotion and Intent Joint Understanding 15967
Yu Zhang, Bin Chen, Hongfei Ye, Zijian Gao, Tianjiao Wan, Long Lan, Kele Xu

Dynamic Graph Convolutional Networks with Spatiotemporal Missing Pattern Awareness 15969
Bingheng Pang, Zhuoxuan Liang, Wei Li, Xiangping Zheng, Rokia Abdein

SIP2Net: Situational-Aware Indoor Pathloss-Map Prediction Network for Radio Map Generation 15974
Wenlihan Lu, Ziyi Lu, Jia Yan, Shijian Gao

Leveraging Multimodal Methods and Spontaneous Speech for Alzheimer's Disease Identification 15976
Yifan Gao, Long Guo, Hong Liu

Accelerometer-Based Person-In-Bed Detection Challenge 15978
Lauren Mentzer, Ravi Kiran Raman, Atulya Yellepeddi

Fast Adaptation of Pretrained Speaker Verification System for Source Speaker Tracking 15980
Xiang Lyu, Yuxuan Wang, Tianyu Zhao, Huadai Liu

TransPathNet: A Novel Two-Stage Framework for Indoor Radio Map Prediction 15982
Xin Li, Ran Liu, Saihua Xu, Sirajudeen Gulam Razul, Chau Yuen

MHAD: Multimodal Home Activity Dataset with Multi-Angle Videos and Synchronized Physiological Signals 15984
Lei Yu, Jintao Fei, Xinyi Liu, Yang Yao, Jun Zhao, Guoxin Wang, Xin Li

Reliable Learning from LLM Features for Multimodal Emotion and Intent Joint Understanding 15989
Xiaolin Xu, Cheng Lu, Zhaoyang Li, Yuyun Liu, Yinghao Ma, Jiahao Luo, Yuan Zong, Wenming Zheng

The USTC System for EEG-Music Emotion Recognition Challenge 15991
Jiaxin Chen, Yiming Wang, Yin-Long Liu, Rui Feng, Jiahong Yuan, Zhen-Hua Ling

Multimodal Fusion for EEG Emotion Recognition in Music with a Multi-Task Learning Framework 15993
Shengyao Huang, Zhishuo Jin, Dongdong Li, Jinchen Han, Xie Tao

Fine-Tuning TitaNet-Large Model for Speaker Anonymization Attacker Systems 15995
Candy Olivia Mawalim, Aulia Adila, Masashi Unoki

Towards Robust Subject Identification from EEG Segments with Data Augmentation Techniques 15997
Meka Nani, Supriyo Banerjea, Raghavan R

Discrete Unit-Based Low-Latency Multi-Lingual Speech Synthesis for LIMMITS'25 Challenge 15999
Yu Jiang, Cheng Gong, Tianrui Wang, Chunyu Qiang, Haoyu Wang, Qiuyu Liu, Yuheng Lu, Xiaobao Wang, Xiaolei Zhang, Longbiao Wang, Jianwu Dang

Cognitive Decline Detection Using DLB Extraction Pipelines 16001
Shibingfeng Zhang, Nadia Khlif, Marcello Ferro, Gloria Gagliardi, Fabio Tamburini

Time-Frequency Based Feature Extraction for Automated Classification of Bed Occupancy Using Accelerometer Data 16003
Akshay Honnavalli, Hrishi Preetham, Gowri Srinivasa

Stream-TTS: A Low-Latency Text-To-Speech Using Kolmogorov-Arnold Networks for Streaming Speech Applications 16006
Giridhar Pamisetty, Riya Ann Easow, Kaustubh Gupta, K Sri Rama Murty

A Multi-Stage Feature Pipeline on Timestamped Speech Transcriptions for Dementia Assessment 16008
Bernhard Thallinger, Laurin Wagner, Theresa Bloder, Mario Zusag

Enhancing Task-Specific Feature Learning with LLMs for Multimodal Emotion and Intent Joint Understanding 16010
Zhaoyang Li, Cheng Lu, Xiaolin Xu, Kaifei Zhang, Yujia Gu, Banghua Li, Yuan Zong, Wenming Zheng

The First VoicePrivacy Attacker Challenge 16012
Natalia Tomashenko, Xiaoxiao Miao, Emmanuel Vincent, Junichi Yamagishi

Multi-Class Dementia Detection Using Acoustic Features - ICASSP-2025 PROCESS Challenge 16014
M. Abdullah Zafar, Xiangyu Zhang, Mostafa Shahin, Beena Ahmed

Person-In-Bed Detection from Mattress-Integrated Accelerometers 16016
Akhil Bhimaraju

The First Indoor Pathloss Radio Map Prediction Challenge 16018
Stefanos Bakirtzis, Çağkan Yapar, Kehai Qiu, Ian Wassell, Jie Zhang

Attacking Voice Anonymization Systems with Augmented Feature and Speaker Identity Difference 16020
Yanzhe Zhang, Zhonghao Bi, Feiyang Xiao, Xuefeng Yang, Qiaoxi Zhu, Jian Guan

Power in Unity: Combining in-Domain and out-Of-Domain Pre-Training Strategies for EEG-Based Person Identification 16022
Christos Garoufis, Marios Glytsos, Ioanna Chourdaki, Panagiotis P. Filntisis, Petros Maragos

TalTech Systems for the PROCESS Signal Processing Grand Challenge 16024
Erik Illaste, Tanel Alumäe

EEG-Music Emotion Recognition: Challenge Overview 16026
Salvatore Calcagno, Simone Carnemolla, Isaak Kavasidis, Simone Palazzo, Daniela Giordano, Concetto Spampinato

Classifying Music-Induced Emotion Using Multi-Modal Ensembles of EEG and Audio Feature Models 16028
Philipp Paukner, Marisa Ripoll, Dilvan Sabir, Deniz Onat Erdoğan, Luca Sacchetto, Klaus Diepold

Early Dementia Detection Using Multiple Spontaneous Speech Prompts: The PROCESS Challenge 16030
Fuxiang Tao, Bahman Mirheidari, Madhurananda Pahar, Sophie Young, Yao Xiao, Hend Elghazaly, Fritz Peters, Caitlin Illingworth, Dorota Braun, Ronan O'Malley, Simon Bell, Daniel Blackburn, Fasih Haider, Saturnino Luz, Heidi Christensen

MEIJU - The 1st Multimodal Emotion and Intent Joint Understanding Challenge 16032
Rui Liu, Xiaofen Xing, Zheng Lian, Haizhou Li, Björn W. Schuller, Haolin Zuo

LIMMITS'25: Multilingual Streaming TTS with Neural Codecs for Indian Languages 16034
Philipp Olbrich, Hema A Murthy, Pranaw Kumar, Shinji Watanabe, Sheng Zhao, Mark Hasegawa-Johnson

Shabdh: A Multi Lingual Zero-Shot Voice Cloning Approach with Speaker Disentanglement 16036
Sreeram Manghat, Sreeja Manghat, Tanja Schultz

MULiving: Towards Real-Time Multi-User Survival State Monitoring Using Wearable RFID Tags..... 16038
Shang Gao, Dawei Yan, Yubo Yan

ScalaLog: Scalable Log-Based Failure Diagnosis Using LLM 16043
Lingzhe Zhang, Tong Jia, Mengxi Jia, Yifan Wu, Hongyi Liu, Ying Li

Aligning Noisy-Clean Speech Pairs at Feature and Embedding Levels for Learning Noise-Invariant Speaker Representations..... 16048
Zuoliang Li, Yang Ai, Jie Zhang, Shengyu Peng, Yu Guan, Bin Gu, Wu Guo

SEF-PNet: Speaker Encoder-Free Personalized Speech Enhancement with Local and Global Contexts Aggregation 16053
Ziling Huang, Haixin Guan, Haoran Wei, Yanhua Long

Improving Open-Vocabulary Video Visual Relation Detection with Decomposed Prompt Learning and Relation Adjustment 16058
Ming Pei, Yi Tan, Yanbin Hao, Hao Zhang, Jinmeng Wu, Basura Fernando, Xun Yang

Spatiotemporal-Aware Visual Captioning Using Vision-Language Pre-Training Model 16063
Shuai Wu, Weidong Yang, Shuyan Wu

Guided Speaker Embedding 16068
Shota Horiguchi, Takafumi Moriya, Atsushi Ando, Takanori Ashihara, Hiroshi Sato, Naohiro Tawara, Marc Delcroix

SiMBA-TS: Simplified Channel Mixing and Mamba for Long-Term Time Series Forecasting..... 16073
Badri Narayana Patro, Vijay Srinivas Agneeswaran

FedImp: Federated Learning Using Important Layers of Client Models for the Diagnosis of Breast Cancer Histopathology Images..... 16078
Mangaldeep Banerjee, Angshuman Paul

DSFormer: Deformable Pointformer for 3D Salient Object Detection 16083
Yibo Hu, Lai Wei, Yanzhe Wang, Qiuyu Huang, Yanding Wei, Qiang Fang

Leave-One-EquiVariant: Alleviating Invariance-Related Information Loss in Contrastive Music Representations 16088
Julien Guinot, Elio Quinton, György Fazekas

IdentityLock: An Identity-Aware Backdoor Strategy for Face Swapping Defense 16093
Zhitao Huang, Peipeng Yu, Zhihua Xia, Xi Yang, Run Lu

NEST: Self-Supervised Fast Conformer as All-Purpose Seasoning to Speech Processing Tasks 16098
He Huang, Taejin Park, Kunal Dhawan, Ivan Medennikov, Krishna C. Puvvada, Nithin Rao Koluguri, Weiqing Wang, Jagadeesh Balam, Boris Ginsburg

Maximizing Connectivity of RIS-Assisted UAV-D2D Networks Using Semidefinite Programming..... 16103
Mohammed Saif, Mohammad Javad-Kalbasi, Shahrokh Valaee

A2B: Neural Rendering of Ambisonic Recordings to Binaural..... 16108
Israel D. Gebru, Todd Keebler, Jake Sandakly, Steven Krenn, Dejan Marković, Julia Buffalini, Samuel Hassel, Alexander Richard

Fast Sparse DFT Computation for Arbitrary Length by Circular Convolution 16113
Soo-Chang Pei, Kuo-Wei Chang

Meta-UAD: A Meta-Learning Scheme for User-Level Network Traffic Anomaly Detection 16118
Tongtong Feng, Qi Qi, Lingqi Guo, Jingyu Wang

Efficient Large-Scale Scene Point Cloud Upsampling with Implicit Neural Networks and Spatial Hashing 16123
Zhiyong Zhang, Yunrui Zhu, Ruyu Liu, Xu Cheng, Jianhua Zhang, Xiufeng Liu

SPRGAN: Streamlined Progressive Refinement for Adversarial Point Cloud Video Upsampling 16128
Zhiyong Zhang, Ruyu Liu, Chaochao Wang, Xianchao Zhang, Jianhua Zhang, Xiufeng Liu

SPEAK: Speech-Driven Pose and Emotion-Adjustable Talking Head Generation 16133
Changpeng Cai, Guinan Guo, Jiao Li, Junhao Su, Fei Shen, Chenghao He, Jing Xiao, Yuanxu Chen, Lei Dai, Feiyu Zhu

Advanced Zero-Shot Text-To-Speech for Background Removal and Preservation with Controllable Masked Speech Prediction 16138
Leying Zhang, Wangyou Zhang, Zhengyang Chen, Yanmin Qian

C2AD: Dual Consistency Learning for Zero-Shot Anomaly Detection 16143
Xiaoling Wang, Ruilong Xing, Zhuotao Tian, Yijun Liu, Senqiao Yang, Yaowei Wang, Jingyong Su

PANDA: Patch-Aware Graph Network with Dual Alignment for Time Series Forecasting 16148
Chen Li, Hongyang Zhang, Saqlain Abbas, Chenyu Ma, Yinhao Liu, Xiaotong Tu

Two-Dimensional Unknown View Tomography from Unknown Angle Distributions 16153
Kaishva Chintan Shah, Karthik S. Gurumoorthy, Ajit Rajwade

Unveiling the Pruning Risks on Privacy Vulnerabilities of Deep Neural Networks 16158
Wenxin Kuang, Qizhuang Liang, Peng Sun, Wei Fu, Qiao Hu, Yupeng Hu

An Automatic Extrinsic Calibration Method for LiDAR-Camera Fusion Via Combining Semantic and Geometric Features 16163
Minqian Wang, Libo Weng, Fei Gao

Toward Comprehensive Semantic Prompt for Region Contrastive Learning Underwater Image Enhancement 16168
Xiao Wang, Yongsheng Fu, Wei Wang, Wei Liu

Multi-Descriptor Mesh Animation Compression 16173
Tai Qin, Chunyang Fu, Ge Li, Shan Liu

MX-Font++: Mixture of Heterogeneous Aggregation Experts for Few-Shot Font Generation 16178
Weihang Wang, Duolin Sun, Jielei Zhang, Longwen Gao

DFNeRF: Disentangled Facial Neural Radiance Fields for Text-Based Editing of Free-View Talking Head 16183
Benwang Chen, Xiaoyu Li, Xuan Wang, Qi Zhang, Haoqian Wang

Adapting Single-Channel Pre-Trained Transformer Models for Multi-Channel Sound Event Localization and Detection 16188
Changjiang He, Siyao Cheng, Jiahua Bao, Jie Liu

LawDNet: Enhanced Audio-Driven Lip Synthesis Via Local Affine Warping Deformation 16193
Junli Deng, Yihao Luo, Xueting Yang, Siyou Li, Wei Wang, Jinyang Guo, Ping Shi

STR-Saliency: Decomposition-Based Perturbations to Generate Saliency Maps for Temporal Black-Box Model Interpretation 16198
Runkai Chen, Yuedi Chen, Haozhe Xu, Xiaopeng Guo, Jun Sun

Text-Aware Adapter for Few-Shot Keyword Spotting 16203
Youngmoon Jung, Jinyoung Lee, Seungjin Lee, Myunghun Jung, Yong-Hyeok Lee, Hoon-Young Cho

Recursive Feature Learning from Pre-Trained Models for Spoofing Speech Detection 16208
Yu Guan, Yang Ai, Zuoliang Li, Shengyu Peng, Wu Guo

AI-Generated Music Detection and Its Challenges 16213
Darius Afchar, Gabriel Meseguer-Brocal, Romain Hennequin

Stochastic Vector Approximate Message Passing with Applications to Phase Retrieval 16218
Hajime Ueda, Shun Katakami, Masato Okada

UPCS: Unbiased Persona Construction for Dialogue Generation 16223
Kuiyun Chen, Yanbin Wei

PIER: A Novel Metric for Evaluating What Matters in Code-Switching 16228
Enes Yavuz Ugan, Ngoc-Quan Pham, Leonard Bärmann, Alex Waibel

A Bayesian Interpretation of Adaptive Low-Rank Adaptation 16233
Haolin Chen, Philip N. Garner

MutualForce: Mutual-Aware Enhancement for 4D Radar-LiDAR 3D Object Detection 16238
Xiangyuan Peng, Huawei Sun, Kay Bierzynski, Anton Fischbacher, Lorenzo Servadei, Robert Wille

Reenvisioning Skeleton-Based Action Recognition Through the Lens of NLP 16243
Long Cao, Shuo Huai, Jingyao Gai

Multimodal Dialogue Emotion Recognition Based on Label Optimization and Coarse-Grained Assisted Fine-Grained 16248
Zhibo Zhang, Zhenyu Yang, Yan Huang, Baojie Xu, Wenyue Hu

Investigating Numerical Translation with Large Language Models 16253
Wei Tang, Jiawei Yu, Yuang Li, Yanqing Zhao, Weidong Zhang, Wei Feng, Min Zhang, Hao Yang

SMCNet: Supervised Surface Material Classification Using mmWave Radar IQ Signals and Complex-Valued CNNs 16258
Stefan Hägele, Fabian Seguel, Driton Salihu, Adam Misik, Eckehard Steinbach

Multi-Microphone Speech Emotion Recognition Using the Hierarchical Token-Semantic Audio Transformer Architecture 16263
Ohad Cohen, Gershon Hazan, Sharon Gannot

You Only Speak Once to See 16268
Wenhao Yang, Jianguo Wei, Wenhuan Lu, Lei Li

DoA-Aided MMSE Channel Estimation for Wireless Communication Systems 16273
Franz Weißer, Nurettin Turan, Wolfgang Utschick

Dual Attention for Space-Time Video Super-Resolution 16278
Jiakai Zheng, Jianping Luo

LLMProto: A Hardware-Efficient Finetuning Model for Few-Shot Relation Extraction with Large Language Model 16283
Longyi Ye, Huaping Zhang

DiffAttack: Imperceptible and Transferable Audio Adversarial Attack Via Diffusion Model 16288
Jiayuan Chen, Yunshu Dai, Fangjun Huang

Deep Enhancement Spotting Network for Low-Complexity Keyword Spotting in Noisy Environments 16293
Yongqiang Chen, Qianhua He, Yanxiong Li, Zunxian Liu, Mingru Yang, Jinxin Huang

FG3DFormer: Fine-Grained 3D Shape Classification Based on Vision Transformer 16298
Xiangyu Ma, Jing Bai, Jinzhe Jiang, Bin Peng

Enhancing Long-Term Capabilities of Large Language Models Via Discourse Sub-Graph Analysis 16303
Zhenyu Guan, Xun Liang, Sensen Zhang

Found in the Distribution: Utilizing Latent Dirichlet Allocation Improves Long Context Comprehension of Large Language Models 16308
Zhenyu Guan, Xun Liang, Sensen Zhang

Learning from Ambiguous Data with Hard Labels 16313
Zeke Xie, Zheng He, Nan Lu, Lichen Bai, Bao Li, Shuo Yang, Mingming Sun, Ping Li

Brick-Diffusion: Generating Long Videos with Brick-To-Wall Denoising 16318
Yunlong Yuan, Yuanfan Guo, Chunwei Wang, Hang Xu, Li Zhang

Codec-ASV: Exploring Neural Audio Codec for Speaker Representation Learning 16323
Yuke Lin, Fulin Zhang, Yingying Gao, Shilei Zhang, Ming Li

Multi-Scale Graph Convolution with Corrective Contrastive Learning for Skeleton-Based Action Recognition 16328
Tianming Zhuang, Erqiang Zhou, Hanwen Zhang, Yi Ding, Ji Geng, Zhen Qin

Relative Localization of Asynchronous Agents Based on Hybrid Active-Passive Two-Way Ranging 16333
Jianing Zhang, Wei Wang, Shiyu Dong, Feng Pan, Baoguo Yu, Xin Li

Multi-Modal Entity Alignment Under Imbalanced Visual Modality Information 16338
Xin Zhang, Yu Liu, Shimin Shan

Improving Multimodal Large Language Models Through Combining Resampler and MLP Projections 16343
Zixing Bai, Yuting Bai

Exploring the Role of CLIP Global Visual Features in Multimodal Large Language Models 16348
Zixing Bai, Yuting Bai

Using Depth-Enhanced Spatial Transformation for Student Gaze Target Estimation in Dual-View Classroom Images 16353
Haonan Miao, Peizheng Zhao, Yuqi Sun, Fang Nan, Xiaolong Zhang, Yaqiang Wu, Feng Tian

MambaNext: An Enhanced Backbone Network with Focus Linear Attention 16358
Dafeng Zhang, Shizhuo Liu

Enhancing Chest X-Ray Classification Through Knowledge Injection in Cross-Modality Learning 16363
Yang Yan, Bingqing Yue, Qiaxuan Li, Man Huang, Jingyu Chen, Zhenzhong Lan

Invariant Model Learning on Local-Aware Wasserstein Geodesic for Domain Adaptation 16368
You-Wei Luo, Yi-Ming Zhai, Chuan-Xian Ren

Author Index