

2024 Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC 2024)

**Macao
3-6 December 2024**

Pages 1-705



**IEEE Catalog Number: CFP2414U-POD
ISBN: 979-8-3503-6734-8**

**Copyright © 2024 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:	CFP2414U-POD
ISBN (Print-On-Demand):	979-8-3503-6734-8
ISBN (Online):	979-8-3503-6733-1
ISSN:	2640-009X

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

CURRAN ASSOCIATES INC.
proceedings
.com

TABLE OF CONTENTS

Multi-Channel Fusion Human Activity Recognition Algorithm Based on Millimeter-Wave Radar	1
<i>Junda Zhu, Shisheng Guo, Longzhen Tang, Guolong Cui</i>	
Run-Time Adaptation of Neural Beamforming for Robust Speech Dereverberation and Denoising	7
<i>Yoto Fujita, Aditya Arie Nugraha, Diego Di Carlo, Yoshiaki Bando, Mathieu Fontaine, Kazuyoshi Yoshii</i>	
Hyperspectral Anomaly Detection Using Robust Principal Component Analysis with Autoencoding Adversarial Network	13
<i>Atsuya Emoto, Ryo Matsuoka</i>	
Data-Driven Tuning for Weighted Least Squares of BLE-AoA-Based Indoor Localization.....	17
<i>Ginji Ohashi, Shinsuke Ibi, Takumi Takahashi, Hisato Iwai</i>	
ET-SSM: Linear-Time Encrypted Traffic Classification Method Based on Structured State Space Model	23
<i>YanJun Li, Xiangyu Zhao, Zhengpeng Zha, Zhenhua Ling</i>	
Iterative Demographic Attentional Feature Fusion-Based CNN and Transformer Network for Accurate Cuffless Blood Pressure Estimation	29
<i>Liwen Tang, Dingchang Zheng, Fei Chen</i>	
Performance Evaluation of Acoustic Echo and Noise Canceller with Variable-Step-Size Shared-Error NLMS Algorithm Under Double-Talk Conditions.....	34
<i>Kenta Iwai, Takanobu Nishiura</i>	
Optimizing Computational Efficiency: In-Memory Computing with Dynamic Switching.....	39
<i>Kun-Lin Tsai, Chao-Ting Huang</i>	
Speech Separation Using Neural Audio Codecs with Embedding Loss.....	45
<i>Jia Qi Yip, Chin Yuen Kwok, Bin Ma, Eng Siong Chng</i>	
Sound Quality Improvement in Visual Microphone by Emphasizing Focused Area Based on Focal Rate.....	51
<i>Hayata Nakano, Yuting Geng, Kenta Iwai, Takanobu Nishiura</i>	
Augmented Sound-Image Perception Using Pre-Virtual-Leading Ultrasounds Based on Precedence Effect	57
<i>Ryota Imanaka, Yuting Geng, Masato Nakayama, Takanobu Nishiura</i>	
Virtual Multi-Boosted Amplitude Modulation Toward High-Pressure Audible Sound with Parametric Array Loudspeakers.....	63
<i>Yoto Ikezaki, Yuting Geng, Masato Nakayama, Takanobu Nishiura</i>	
Deep-Learning-Based Speech Enhancement with Rough-Focused Optical Laser Microphone by Reconstructing Complex Spectrum	69
<i>Yuki Nakano, Yuting Geng, Kenta Iwai, Takanobu Nishiura</i>	
A Study on Multimodal Fusion and Layer Adapter in Emotion Recognition.....	74
<i>Xiaohan Shi, Yuan Gao, Jiajun He, Jinyi Mi, Xingfeng Li, Tomoki Toda</i>	
Analyzing House Music: Relations of Audio Features and Musical Structure.....	80
<i>Justin Tomoya Wulf, Tetsuro Kitahara</i>	

Band-Split Inter-SubNet: Band-Split with Subband Interaction for Monaural Speech Enhancement.....	85
<i>Yen-Chou Pan, Yih-Liang Shen, Yuan-Fu Liao, Tai-Shih Chi</i>	
Speech Synthesis from IPA Sequences Through EMA Data	91
<i>Koki Maruyama, Shun Sawada, Hidefumi Ohmura, Kouichi Katsurada</i>	
Enhancing Shear Wave Propagation Analysis in Tissue with Directional Filtering of Reflected Waves	97
<i>Quang Hai Luong, Duc-Nghia Tran, Nguyen Sy Hiep, Lam Sinh Cong, Duc-Tan Tran</i>	
A Survey on Objective Quality Assessment of Omnidirectional Images	103
<i>Xiangjie Sui, Shiqi Wang, Yuming Fang</i>	
Enhancing Security Using Random Binary Weights in Privacy-Preserving Federated Learning.....	109
<i>Hiroto Sawada, Shoko Imaizumi, Hitoshi Kiya</i>	
Modeling and Analysis of the Interaction Between Opinions and Actions Among Heterogeneous Agents.....	115
<i>Hangjing Zhang, H. Vicky Zhao</i>	
ViP-CBM: Reducing Parameters in Concept Bottleneck Models by Visual-Projected Embeddings	121
<i>Ji Qi, Huisheng Wang, H. Vicky Zhao</i>	
BEES: A New Acoustic Task for Blended Emotion Estimation in Speech.....	127
<i>Xingfeng Li, Xiaohan Shi, Yuke Si, Zilong Zhang, Feifei Cui, Yongwei Li, Yang Liu, Masashi Unoki, Masato Akagi</i>	
Successive Speaker Relative Transfer Function Estimation Through Relative Transfer Matrix in Noisy Reverberant Environments.....	133
<i>Wageesha N. Manamperi, Thushara D. Abhayapala</i>	
Relative Transfer Matrix for Drone Audition Applications: Source Enhancement	139
<i>Wageesha N. Manamperi, Thushara D. Abhayapala</i>	
Sampling Pattern Augmentation to Enhance Deep Learning-Based Image Reconstruction of MRI.....	145
<i>Kazuki Yamato, Satoshi Ito</i>	
Enhancing YOLOv7 with GLF-Trans for Precision in Small Object Detection.....	151
<i>Naohito Yoshikawa, Masaaki Ikehara</i>	
Structural Analysis of Asian and African Rice Panicles Via Transfer Learning	156
<i>Cong Hieu Le, Lam Thai Nguyen, Trung Kien Pham, Le Khanh Nguyen, Tran Hiep Dinh, Stefan Jouannic, Helene Adam, Pierre Duhammel, Nguyen Linh Trung, Hoang Trong Minh</i>	
Adaptive Subspace Clustering for Matrix Completion.....	164
<i>Takuto Wada, Ryohei Sasaki, Katsumi Konishi</i>	
New Approach for Alzheimer's Disease Classification Using Topographic Maps and Deep Learning Model	169
<i>Quoc Anh Le, Hong Thinh Nguyen</i>	
M-IRRA: A Multilingual Model for Text-Based Person Search.....	175
<i>Hung-Phong Tran, Thi-Hoai Phan, Thuy-Binh Nguyen, Thi-Ngoc-Diep Do, Hong-Quan Nguyen, Thanh-Hai Tran, Hien-Thanh Duong, Thi-Lan Le</i>	

Data Augmentation and Assessment for Enhanced Ovarian Tumor Classification	181
<i>Thi-Loan Pham, Gia-Minh Pham, Tien-Dat Nguyen, Van-Hung Le, Thi-Lan Le, Duy-Hai Vu, Hai Vu, Chi-Mai Pham, Thanh-Hai Tran</i>	
Ablation Study to Derive a Computationally Efficient Deep Learning-Based Super-Resolution Approach	187
<i>Asfa Jamil, Alessandro Artusi</i>	
Postoperative Delirium Prediction Based on Preoperative Electrocardiogram and Electroencephalogram	193
<i>Shogo Mito, Miho Miyajima, Hirofumi Tomioka, Hitomi Sato, Takashi Takeuchi, Hitoshi Muto, Yuji Kabasawa, Hiroyuki Harada, Kana Eguchi, Shota Kato, Manabu Kano</i>	
Observation of the Terrestrial Radio Environment Using the Low Earth Orbit Satellite Constellation	198
<i>Takatoshi Obata, Osamu Takyu, Kei Inage, Takeo Fujii, Kohei Yoshida, Masayuki Ariyoshi</i>	
A High-Isolation Sub-6 GHz In-Band Full-Duplex Communication System	203
<i>Chengzhe Shi, Wensheng Pan, Wanzhi Ma, Ying Liu, Qiang Xu, Zhiya Zhang, Shihai Shao</i>	
A Parameter-Free Model for Long-Term Concrete Creep Prediction	209
<i>Conghui Li, Chern Hong Lim, Xin Wang</i>	
Learnable Cross-Correlation Based Filter-And-Sum Networks for Multi-Channel Speech Separation	215
<i>Xianrui Wang, Shiqi Zhang, Bo He, Shoji Makino, Jingdong Chen</i>	
Is Corpus Suitable for Human Perception?: Quality Assessment of Voice Response Timing in Conversational Corpus Through Timing Replacement	220
<i>Sadahiro Yoshikawa, Ryo Ishii, Shogo Okada</i>	
Enhancing Branchformer with Dynamic Branch Merging Module for Code-Switching Speech Recognition	226
<i>Hong-Jie Hu, Yu-Chiao Lai, Chia-Ping Chen</i>	
A Solution for Anomaly Detection of Red Beans in a Product Processing Line	232
<i>Duc Hai Nguyen, Trong Hiep Do, Hoang Linh Phuong Nguyen, Quoc Khanh Nguyen, Duc Tan Tran, Tien-Son Bui, Van Toi Nguyen</i>	
An Isolated Vietnamese Sign Language Recognition Method Using a Fusion of Heatmap and Depth Information Based on Convolutional Neural Networks	237
<i>Xuan-Phuoc Nguyen, Thi-Huong Nguyen, Duc-Tan Tran, Tien-Son Bui, Van-Toi Nguyen</i>	
Enhancing Neural Speech Embeddings for Generative Speech Models	243
<i>Doyeon Kim, Yanjue Song, Nilesch Madhu, Hong-Goo Kang</i>	
Heavy-Tailed Distributions-Based Online Semi-Blind Source Separation for Nonlinear Echo Cancellation	249
<i>Liyan Zhang, Xianrui Wang, Yichen Yang, Tetsuya Ueda, Shoji Makino, Jingdong Chen</i>	
GILED: Lesion Detection of Gastrointestinal Tract from Endoscopic Images and Medical Notes	254
<i>Vu-An Hoang, Minh-Hanh Tran, Viet Hang Dao, Thanh-Hai Tran</i>	
Prediction-Error-Based Adaptive SpecAugment for Fine-Tuning the Masked Model on Audio Classification Tasks	260
<i>Xiao Zhang, Haoran Xing, Mingxue Song, Daiki Takeuchi, Noboru Harada, Shoji Makino</i>	

Generalized Graph Signal Sampling Under Subspace Priors by Difference-Of-Convex Minimization	266
<i>Keitaro Yamashita, Kazuki Naganuma, Shunsuke Ono</i>	
Gamma-VAE: Speech Representation Based on VAE Assuming Gamma Distribution for Both Latent Variables and Observation	272
<i>Nanako Imaichi, Toru Nakashika</i>	
A Method for Classification NEO–FFI Answers Fabricated and Advantageous Due to Psychological Bias Using Brainwave Specific Brain Activity Networks	278
<i>Yuto Ashikawa, Takashi Ito, Shohei Ishizu, Yosuke Kurihara</i>	
Optimizing Multi-Speaker Speech Recognition with Online Decoding and Data Augmentation	282
<i>Yizhou Peng, Eng Siong Chng</i>	
Adapting OpenAI’s Whisper for Speech Recognition on Code-Switch Mandarin-English SEAME and ASRU2019 Datasets	288
<i>Yuhang Yang, Yizhou Peng, Hao Huang, Eng Siong Chng, Xionghu Zhong</i>	
Does Brain Atlas Choice Matter? an Empirical Study in Alzheimer’s Diagnosis Using FDG-PET Images	294
<i>Pham Minh Tuan, Mouloud Adel, Nguyen Linh Trung, Eric Guedj</i>	
A Configurable OFDM Baseband Processor for RF-UOWC System-On-Chip	300
<i>Trio Adiono, Erwin Setiawan, Michael Jonathan, Rahmat Mulyawan, Nana Sutisna, Infall Syafalni, Wasiiu O. Popoola</i>	
Monocular Depth Estimation for Autonomous Driving Based on Instance Clustering Guidance	304
<i>Dahyun Kim, Dongkwon Jin, Chang-Su Kim</i>	
Transformer Attention Matrix Multiplication Design Using 4×4 Systolic Arrays	310
<i>Muhammad Sayyid Afif, Infall Syafalni, Nana Sutisna, Trio Adiono</i>	
A Single-Input/Binaural-Output Perceptual Rendering Based Speech Separation Method in Noisy Environments	316
<i>Tianqin Zheng, Hanchen Pei, Ningning Pan, Jilu Jin, Gongping Huang, Jingdong Chen, Jacob Benesty</i>	
GMA: Green Multi-Modal Alignment for Image-Text Retrieval	321
<i>Tsung-Shan Yang, Yun-Cheng Wang, Chengwei Wei, Suya You, C.-C. Jay Kuo</i>	
Affine Combination of General Adaptive Filters	327
<i>Danqi Jin, Yitong Chen, Jie Chen, Gongping Huang</i>	
Psychological Driving Style Estimation from GPS Sensor Data Alone	332
<i>Hiroto Horimoto, Ryusei Kimura, Takahiro Tanaka, Shogo Okada</i>	
Physical Domain Adversarial Attacks Against Source Printer Image Attribution	338
<i>Nischay Purnekar, Benedetta Tondi, Mauro Barni</i>	
An Annealing-Inspired Gradient-Descent Based Suboptimal Solver for Combinatorial Problems	344
<i>Shu-Ping Chang, Cheng-Che Lee, Hsin-Jung Lee, Chieh-Hsiung Kuan, Jason Gemsun Young, Chia-Yu Yao, Jian-Jiun Ding</i>	
Optimising Neural Networks with Fine-Grained Forward-Forward Algorithm: A Novel Backpropagation-Free Training Algorithm	350
<i>James Gong, Bruce Li, Waleed Abdulla</i>	

A Noisy Context Optimization Approach for Chinese Spelling Correction	356
<i>Guangwei Zhang, Yongping Xiong, Ruifan Li</i>	
Quefrency Approach to Audio Deepfake Detection	362
<i>Kanishq Singhal, Aditya Goyal, Priyanka Gupta</i>	
Effect of Dynamic Binaural Beats on Concentration Enhancement	368
<i>Jun-Seok Lee, Yun-Sung Lee, Han-Jeong Hwang</i>	
Beamforming Informed Independent Low-Rank Matrix Analysis for Sound Source Enhancement in Unmanned Aerial Vehicles	372
<i>Jin Xuan Teh, Norihiro Takamune, Hiroshi Saruwatari, Benjamin Yen, Michael Kingan, Yusuke Hioka</i>	
Wavetable Synthesis Using CVAE for Timbre Control Based on Semantic Label.....	378
<i>Tsugumasa Yutani, Yuya Yamamoto, Shuyo Nakatani, Hiroko Terasawa</i>	
Frequency & Channel Attention Network for Small Footprint Noisy Spoken Keyword Spotting	384
<i>Yuanxi Lin, Yuriy Evgenyevich Gapanyuk</i>	
A Diffusion-Based Approach for Restoring Face-Swapped Images.....	390
<i>Yuanchen Niu, Yuanman Li, Guijia Zhang, Xia Li</i>	
A Semi-Supervised Low-Light Image Enhancement with Color Guidance	395
<i>Yuxin Wang, Shuolin Yang, Qianxi Wu, Zhishuo Zhang, Yunxia Liu, Yang Yang, Yakui Dong, Cheng Fei, Junliang Liu, Lili Wang, Shuzhen Fan, Yongfu Li</i>	
Robust Watermarking Via Dual Guidance.....	401
<i>Yuhang Zhang, Yuanman Li, Li Dong, Xia Li</i>	
Long Audio File Speaker Diarization with Feasible End-To-End Models	407
<i>Kai-Wei Huang, Chia-Ping Chen</i>	
Rotation Invariant Spatio-Spectral Total Variation for Hyperspectral Image Denoising.....	413
<i>Shingo Takemoto, Shunsuke Ono</i>	
Hyperspectral Unmixing with Row-Sparsity Enhancement: A Difference-Of-Convex Approach	419
<i>Kazuki Naganuma, Shunsuke Ono</i>	
Cloud Removal in Hyperspectral Satellite Images Using Low-Rank Tensor Completion	424
<i>Chuong Hoang Vo, Truong Thanh Nhat Mai, Chul Lee</i>	
Region Aware Framework for Constrained Image Splicing Detection and Localization	430
<i>Haokun Cao, Yuanman Li, Xinyu Yang, Xia Li</i>	
Estimation of Rotation Angle and Anisotropic Scaling Rate Using Pilot Signals for Watermarking	436
<i>Rinka Kawano, Masaki Kawamura</i>	
Deep Unfolding Aided Parameter Optimization for Multi-Task Diffusion LMS Algorithm.....	442
<i>Xiaoqing Tong, Kazunori Hayashi</i>	
Analysis of Various Self-Supervised Learning Models for Automatic Pronunciation Assessment	448
<i>Haeyoung Lee, Sunhee Kim, Minhwa Chung</i>	
On the Security of Bitstream-Level JPEG Encryption with Restart Markers.....	454
<i>Mare Hirose, Shoko Imaizumi, Hitoshi Kiya</i>	

Design of Spectrogram-Consistency Regularization Term Dependent on Observation in Independent Low-Rank Matrix Analysis for Blind Source Separation	460
<i>Takaaki Kojima, Norihiro Takamune, Daichi Kitamura, Hiroshi Saruwatari</i>	
Teager Energy Cepstral Coefficients for Spoken Language Identification.....	466
<i>Arth J. Shah, Savita H. Yadav, Hemant A. Patil</i>	
Two-Way Malaysian Sign Language Communication System for Inclusive Education	472
<i>Veron Zhen Liang Hii, Aaron Ken Kiat Lo, Ida Pei Xin Lee, Alec Vince Gonzales Abuan, Sue Han Lee, Patrick Hang Hui Then</i>	
Voice Liveness Detection Using Linear Frequency Residual Cepstral Coefficients	478
<i>Arth J. Shah, Nandini V. Mandaviya, Hemant A. Patil</i>	
Improved Ultimate Link Without Markers for Projective Transformation.....	484
<i>Keiji Yamadera, Michiharu Niimi</i>	
Detection of Diffusion-Generated Images Using Sparse Coding	490
<i>Daishi Tanaka, Michiharu Niimi</i>	
Adversarial Augmentation and Adaptation for Speech Recognition	496
<i>Jen-Tzung Chien, Wei-Yu Sun</i>	
Generation of Target Speech with Speaker Individuality Based on Accent Conversion for English Pronunciation Learning	502
<i>Rei Hamakawa, Michiharu Niimi</i>	
Empathetic Response Generation Via Regularized Q-Learning.....	508
<i>Jen-Tzung Chien, Yi-Chien Wu</i>	
On Joint Dereverberation and Single Moving Source Separation with Online Source Steering	514
<i>Yiting Zhang, Kaien Mo, Tetsuya Ueda, Yichen Yang, Shoji Makino</i>	
Block Refinement Learning for Improving Early Exit in Autoregressive ASR.....	518
<i>Naotaka Kawata, Shota Orihashi, Satoshi Suzuki, Tomohiro Tanaka, Mana Ihori, Naoki Maikishima, Taiga Yamane, Ryo Masumura</i>	
Contrastive Learning Based Knowledge Distillation for Enhancing Defect Detection.....	524
<i>Jing-Ming Guo, Lun-Da Yuan, Cian Huang, Yi-Chong Zeng</i>	
Color Guided Disease Segmentation for Plant Images.....	530
<i>So-Yeon Jang, Jong-Ok Kim</i>	
A Novel Kind of WVD Associated with the Linear Canonical Transform.....	536
<i>Jia-Yin Peng, Jian-Yi Chen, Bing-Zhao Li</i>	
New Perspectives and Insights on Distortionless Microphone Array Beamforming.....	542
<i>Fan Zhang, Jacob Benesty, Chao Pan, Jingdong Chen</i>	
How Accurate Can Large Vision Language Model Perform for Images with Compression Degradation?	547
<i>Xiaohan Fang, Peilin Chen, Meng Wang, Shiqi Wang</i>	
A Quasilinear-Time CVP Algorithm for Triangular Lattice Based Fuzzy Extractors and Fuzzy Signatures	553
<i>Kenta Takahashi, Wataru Nakamura</i>	

Effect of White Noise on Working Memory Using Event-Related Potentials	557
<i>Seung-Won Lee, Jun-Seok Lee, Han-Jeong Hwang</i>	
Performance Optimization in the Cascade of VAD and ASR Systems: A Study on Evaluation and Alignment Strategies	561
<i>Zhentao Lin, Zihao Chen, Bi Zeng, Leqi Chen, Jia Cai</i>	
ComplexFace: A Public Visible-Thermal Face Dataset with Real-Life Complexity	567
<i>Jiajin He, Chengxi Dong, Yunqi Cai, Dong Wang</i>	
A Discrete-Valued Signal Estimation by Nonconvex Enhancement of SOAV with cLiGME Model	573
<i>Satoshi Shoji, Wataru Yata, Keita Kume, Isao Yamada</i>	
U-Mamba-Net: A Highly Efficient Mamba-Based U-Net Style Network for Noisy and Reverberant Speech Separation	579
<i>Shaoliang Dang, Tetsuya Matsumoto, Yoshinori Takeuchi, Hiroaki Kudo</i>	
Proposal of Blind Extractable Additive Video Watermarking Method	584
<i>Nao Harada, Rinka Kawano, Masaki Kawamura</i>	
StylebookTTS: Zero-Shot Text-To-Speech Leveraging Unsupervised Style Representation	590
<i>Juhwan Yoon, Hyungseob Lim, Hyeonjin Cha, Hong-Goo Kang</i>	
Knowledge Augmented Attention Gating Embedding for Link Prediction	596
<i>Zewei Chen, Shuhong Chen, Chen Li, Xianwei Zheng, Minfan He, Xutao Li</i>	
Low-Complexity Adaptive Beamformer for Joint Reverberation and Noise Suppression	601
<i>Fan Zhang, Chao Pan, Jingdong Chen, Jacob Benesty</i>	
Graph Filter Transfer for Time-Varying Signal Estimation Between Two Networks	606
<i>Tsutahiro Fukuhara, Junya Hara, Hiroshi Higashi, Yuichi Tanaka</i>	
Transfer-Based Adversarial Attack Against Multimodal Models by Exploiting Perturbed Attention Region	612
<i>Raffaele Disabato, Aprilpyone Maungmaung, Huy H. Nguyen, Isao Echizen</i>	
Multichannel Speech Enhancement Using Complex-Valued Graph Convolutional Networks and Triple-Path Attentive Recurrent Networks	618
<i>Xingyu Shen, Wei-Ping Zhu</i>	
Speech Dereverberation with Deconvolution Regularized by Denoising	624
<i>Haonan Hu, Ziye Yang, Jie Chen, Lijun Zhang</i>	
Enhanced RefineDNet for Single Image Dehazing	630
<i>Jingyu Ren, Lei Yang</i>	
Domain Adaptation by Alternating Learning of Acoustic and Linguistic Information for Japanese Deaf and Hard-Of-Hearing People	636
<i>Kaito Takahashi, Yukoh Wakabayashi, Kengo Ohta, Akio Kobayashi, Norihide Kitaoka</i>	
YOLO for High Resolution Images Without Retraining	643
<i>Daisuke Minami, Kiyoshi Nishikawa</i>	
Toward Universal Detector for Synthesized Images by Estimating Generative AI Models	649
<i>Ryota Seo, Minoru Kuribayashi, Akinobu Ura, Antoine Mallet, Rémi Cogranne, Wojciech Mazurczyk, David Megías</i>	

Anomalous Machine Sound Detection Based on Time Domain Gammatone Spectrogram Feature and IDNN Model.....	655
<i>Primanda Adyatma Hafiz, Candy Olivia Mawalim, Dessi Puji Lestari, Sakriani Sakti, Masashi Unoki</i>	
Speech Emotion Recognition Based on Crossmodal Transformer and Attention Weight Correction	661
<i>Ryusei Terui, Takeshi Yamada</i>	
Adaptive Time-Varying Graph Learning for Traffic Flow Data Based on Anomaly Moment Detection	666
<i>Shuhong Chen, Zewei Chen, Chen Li, Xianwei Zheng, Minfan He, Xutao Li</i>	
Unsupervised Anomalous Sound Detection Using Timbral and Human Voice Disorder-Related Acoustic Features	671
<i>Malik Akbar Hashemi Rafsanjani, Candy Olivia Mawalim, Dessi Puji Lestari, Sakriani Sakti, Masashi Unoki</i>	
Tsnake: A Time-Embedded Recurrent Contour-Based Instance Segmentation Model.....	677
<i>Chen-Jui Hsu, Jian-Jiun Ding, Chun-Jen Shih</i>	
Generating Phonetic Transcriptions for Korean English L2 Learners Using Multiple Self-Supervised-Model-Based ASR Systems and Rover Method	683
<i>Jong In Kim, Sunhee Kim, Minhwa Chung</i>	
Unsupervised Discovery of Non-Categorical L2 Error Patterns Using Wav2Vec2.0 Code Vectors.....	689
<i>Eunsoo Hong, Sunhee Kim, Minwha Chung</i>	
Speech Depression Recognition from the Selfreference Effect Using LSTM with ResNet	695
<i>Daimin Shi, Xiaoyong Lu, Yang Liu, Jingyi Yuan, Tao Pan</i>	
A Permutation-Based Reversible Data Hiding Method with Zero Visual Distortion	700
<i>Wendi Zhu, Koksheik Wong, Minoru Kuribayashi</i>	
Improved Architecture for High-Resolution Piano Transcription to Efficiently Capture Acoustic Characteristics of Music Signals	706
<i>Jinyi Mi, Sehun Kim, Tomoki Toda</i>	
Ev3DGS: Event Enhanced 3D Gaussian Splatting from Blurry Images	712
<i>Junwu Huang, Zhexiong Wan, Zhicheng Lu, Juanjuan Zhu, Mingyi He, Yuchao Dai</i>	
PRTGaussian: Efficient Relighting Using 3D Gaussians with Precomputed Radiance Transfer	718
<i>Libo Zhang, Yuxuan Han, Wenbin Lin, Jingwang Ling, Feng Xu</i>	
Automated Prediction of Loudness Growth Curve Using EEG Signals.....	724
<i>Nitya Tiwari, Arjun Reddy Vadyala, K. S. Nataraj</i>	
EEG-Based Evaluation of Enjoyment Emotion During Cognitive-Motor Task	730
<i>Haruna Aoki, Sinan Zhang, Yumie Ono</i>	
Generation of Photo Slideshow with Song Based on Closeness Between Concept of Lyrics and that of Images	734
<i>Mei Hashimoto, Michiharu Niimi</i>	
Screen Content Encoding Network Based on Deep Contextual Information	740
<i>Tianyu Gong, Tao Zhang, Ye Zhong, Mengmeng Zhang, Huihui Bai</i>	

Exploring Brain Connectivity Patterns and Cognitive Resilience in Aging: A Study with the LEMON Dataset.....	746
<i>Kapeleshh Ks, Wei Chen, Prince Aldrin Domer, Hong Ji</i>	
SDNet: Noise-Robust Bandwidth Extension Under Flexible Sampling Rates.....	752
<i>Junkang Yang, Hongqing Liu, Lu Gan, Yi Zhou, Xing Li, Jie Jia, Jinzhuo Yao</i>	
Few-Shot Audio Classification Model for Detecting Classroom Interactions Using LaSO Features in Prototypical Networks.....	758
<i>Rashed Iqbal, Christian Ritz, Jack Yang, Sarah Howard</i>	
Innovative Information Hiding in H.266/VVC Using Sub-Block Transform Technique	764
<i>Hau Joan, Yiqi Tew, Li Peng Tan</i>	
Cuisine Image Synthesis with Improved Multiscale GANs Guided by CLIP	770
<i>Weiyi Xia, Satoru Fujita</i>	
Disposable-Key-Based Image Encryption for Collaborative Learning of Vision Transformer	776
<i>Rei Aso, Sakaya Shiota, Hitoshi Kiya</i>	
A Novel LLM-Based Two-Stage Summarization Approach for Long Dialogues	782
<i>Yuan-Jhe Yin, Bo-Yu Chen, Berlin Chen</i>	
Impulse Response Transforming Method to Control Distance Perception Based on Direct-To-Reverberant Energy Ratio	788
<i>Toru Takahashi, Eita Morigaki, Masato Nakayama</i>	
Significance of Lower Frequency Regions for Audio Deepfake Detection.....	794
<i>Arth J. Shah, Hemant A. Patil</i>	
New Abnormal Behavior Detection for Patient Surveillance System	800
<i>Yujin Han, Taewan Kim</i>	
Effect of Phase-Locked Transcranial Alternating Current Stimulation on Vocal Tremor.....	805
<i>Junting Wang, Satoko Koganemaru, Atsushi Shima, Yedi Cao, Kana Hirakawa, Ken Iwagana, Atsushi Suehiro, Keiko Maekawa, Tatsuya Mima, Yumie Ono</i>	
Utilizing Cross Layer Attentions for Semantic Segmentation of Small Objects	811
<i>Yu-Hsien Chung, Chi-Hsuan Lu, Jung-Hui Cho, Chih-Chang Yu</i>	
Data-Driven Sound Field Reproduction for Higher-Order Mode Matching Using a Circular Loudspeaker Array	817
<i>Keiko Kawase, Gen Sato, Izumi Tsunokuni, Yusuke Ikeda</i>	
Subset Random Sampling of Finite Time-Vertex Graph Signals.....	822
<i>Hang Sheng, Qinji Shu, Hui Feng, Bo Hu</i>	
An Effective Contextualized Automatic Speech Recognition Approach Leveraging Self-Supervised Phoneme Features.....	828
<i>Li-Ting Pai, Yi-Cheng Wang, Bi-Cheng Yan, Hsin-Wei Wang, Jia-Liang Lu, Chi-Han Lin, Juan-Wei Xu, Berlin Chen</i>	
Real-Time Monophonic Dual-Pitch Extraction Model.....	834
<i>Ngoc-Son Tran, Pei-Chin Hsieh, Yih-Liang Shen, Yen-Hsun Chu, Tai-Shi Chi</i>	

SMoLnet-T: An Efficient Complex-Spectral Mapping Speech Enhancement Approach with Frame-Wise CNN and Spectral Combination Transformer for Drone Audition	840
<i>Zhi-Wei Tan, Andy W. H. Khong</i>	
Reduced-Dimensional MUSIC Algorithm for Frequency Diverse Array in MIMO Radar System	846
<i>Beizuo Zhu, Kazunori Hayashi, Hiroki Mori</i>	
A Coarse-To-Fine Change Detection Method for Remote Sensing Sparse Cultivated Land.....	854
<i>Yuan Hu, Yifan Zhang, Mingyang Ma, Shaohui Mei</i>	
RepViT Based Lightweight Architecture for Distracted Driving Detection.....	860
<i>Muwei Jian, Yukun Ling, Rui Wang, Yanjie Zhong, Huihui Huang, Xiaoguang Li</i>	
Layer-Wise Feature Distillation with Unsupervised Multi-Aspect Optimization for Improved Automatic Speech Assessment	866
<i>Chung-Wen Wu, Berlin Chen</i>	
Wind Noise Reduction with Orthogonal Polynomial Expansion	871
<i>Li Du, Chao Pan, Lijun Zhang</i>	
Sparse Blind Deconvolution and Demixing Via Block Majorization-Minimization.....	876
<i>Mengting Chen, Ziping Zhao</i>	
An End-To-End Two-Stream Network Based on RGB Flow and Representation Flow for Human Action Recognition.....	882
<i>Song-Jiang Lai, Tsun-Hin Cheung, Ka-Chun Fung, Tian-Shan Liu, Kin-Man Lam</i>	
Few-Shot Open-Set Keyword Spotting with Multi-Stage Training.....	888
<i>Lo-Ya Li, Tien-Hong Lo, Jieih-Weih Hung, Shih-Chieh Huang, Berlin Chen</i>	
GLASS: Investigating Global and Local Context Awareness in Speech Separation.....	893
<i>Kuan-Hsun Ho, En-Lun Yu, Jieih-Weih Hung, Shih-Chieh Huang, Berlin Chen</i>	
Self-Supervised Augmented Diffusion Model for Anomalous Sound Detection	899
<i>Jiawei Yin, Wenbin Zhang, Mingjun Zhang, Yu Gao</i>	
Detecting Abnormal Machine Sounds Using an Ensemble Approach with Data Augmentation Techniques.....	904
<i>Po Cheng Chan, Chung Li Lu, Jia Ching Wang</i>	
Integrating VGGSK and BEATs for Enhanced Sound Event Detection: A Semi-Supervised GRU-Based System with Weak Labels and Synthetic Soundscapes.....	908
<i>Po Cheng Chan, Wei-Yu Chen, Chung Li Lu, Hsiang-Feng Chuang, Yu-Han Cheng, Jia Ching Wang</i>	
Drone Audition: Implementation of an Indoor Multi-Drone System for Sound Source Tracking.....	913
<i>Benjamin Yen, Kazuhiro Nakadai</i>	
Music2Fail: Transfer Music to Failed Recorder Style.....	919
<i>Chon In Leong, I-Ling Chung, Kin-Fong Chao, Jun-You Wang, Yi-Hsuan Yang, Jyh-Shing Roger Jang</i>	
Murmur Separation and Classification from Heart Sound Using Constrained Singular Spectrum Analysis and Wavelet Transform.....	924
<i>Yuanyang Qi, Saeid Sanei</i>	

VietSing: A High-Quality Vietnamese Singing Voice Corpus	929
<i>Minh Vu, Zhou Wei, Binit Bhattacharai, Kah Kuan Teh, Tran Huy Dat</i>	
Learning a Sequence of Cursive-Style Japanese Characters in Classical Literary Works	935
<i>Satoru Fujita, Keizo Oyama</i>	
Inertial Strengthened CLIP Model for Zero-Shot Multimodal Egocentric Activity Recognition.....	941
<i>Mingzhou He, Haojie Wang, Shuchang Zhou, Qingbo Wu, King Nghi Ngan, Fanman Meng, Hongliang Li</i>	
COIN-AT-PVAD: A Conditional Intermediate Attention PVAD	947
<i>En-Lun Yu, Ruei-Xian Chang, Jieh-Weih Hung, Shih-Chieh Huang, Berlin Chen</i>	
Separation of Cardiopulmonary Sound Signals for Classification of Respiratory Diseases.....	952
<i>Ruxin Zheng, Saeid Sanei</i>	
Collection of Correlated Information from Superimposed Multiple Chirp Signals.....	958
<i>Koki Aoyama, Koichi Adachi</i>	
Hammering Inspection System Using HPSS and Gradient Boosting with a Wall-Climbing Robot.....	964
<i>Nichika Koyama, Nari Tanabe, Masaya Fujisawa</i>	
Optimization of the Intensity Aware Loss for Dynamic Facial Expression Recognition.....	969
<i>Davy Tec-Hinh Lau, Jian-Jiun Ding, Guillaume Muller</i>	
Dynamic Sensor Placement on Graphs Based on Graph Signal Sampling Theory	974
<i>Saki Nomura, Junya Hara, Hiroshi Higashi, Yuichi Tanaka</i>	
Automatic Prompt Generation and Grounding Object Detection for Zero-Shot Image Anomaly Detection	980
<i>Tsun-Hin Cheung, Ka-Chun Fung, Songjiang Lai, Kwan-Ho Lin, Vincent Ng, Kin-Man Lam</i>	
A Non-Intrusive Speech Quality Assessment Model Using Whisper and Multi-Head Attention.....	986
<i>Guojian Lin, Yu Tsao, Fei Chen</i>	
Implementation of Real-Time Oscillometric Based Algorithm for Blood Pressure Measurement in Patient Monitor.....	992
<i>Trio Adiono, Clarence Amadeus, Sindy Novaria Cicilya Sinaga, Teuku Rafisyah Thomi</i>	
Exploration Robot Based on YOLOv8 Algorithm.....	998
<i>Infal Syafalni, Angelica Winasta Sinisuka, Dwi Kalam Amal Tauhid, Farrel Ahmad, Muhammad Alif Putra Yasa, Steven Alexander Wen, Erwin Setiawan, Nana Sutisna, Trio Adiono</i>	
Development of Simple Algorithm to Detect and Filter Motion Artifact Noise in Non Invasive Blood Pressure (NIBP) Measurement	1003
<i>Trio Adiono, Rd Elviana La'Salina Muhlis, Clarence Amadeus, Sindy Novaria Cicilya Sinaga</i>	
AI-Generated Image Detectors Are Surprisingly Easy to Mislead... for Now.....	1009
<i>Zihang Lyu, Jun Xiao, Cong Zhang, Kin-Man Lam</i>	
Multibiometrics Using a Single Face Image	1014
<i>Koichi Ito, Taito Tonosaki, Takafumi Aoki, Tetsushi Ohki, Masakatsu Nishigaki</i>	
Dictionary Learning for Directed Graph Signals Via Augmented GFT	1020
<i>Tsubasa Naito, Ryuto Ito, Yuichi Tanaka, Shogo Muramatsu</i>	

MYMV: A Music Video Generation System with User-Preferred Interaction.....	1026
<i>Kyungjune Lee, Mingyu Jang, Jungwoo Huh, Jeonghaeng Lee, Seokkeun Choi, Sanghoon Lee</i>	
Temporal-Spatial Correlation Analysis for Ship-Radiated Noise Based on Random Matrix Theory.....	1030
<i>Qing Feng, Zhiqiang Wu, Xuebin Li, Heping Shen, Liushang, Tangmin, Shengquan Feng</i>	
Can We Estimate Purchase Intention Based on Zero-Shot Speech Emotion Recognition?.....	1036
<i>Ryotaro Nagase, Takashi Sumiyoshi, Natsuo Yamashita, Kota Dohi, Yohei Kawaguchi</i>	
Text-Guided Visual Prompt Tuning with Masked Images for Facial Expression Recognition	1042
<i>Rongkang Dong, Cuixin Yang, Kin-Man Lam</i>	
Noise-Robust Estimation of Early-Part Room Impulse Responses Based on Physics-Informed Neural Network with Dynamic Pulling Method.....	1048
<i>Ken Kurata, Gen Sato, Izumi Tsunokuni, Yusuke Ikeda</i>	
Fine-Grained Privacy-Preserving Image Retrieval in Cloud Environment	1053
<i>Jing Liang, Libo Wang, Peiya Li</i>	
Assessment and Improvement of Customer Service Speech with Multiple Large Language Models	1059
<i>So Watanabe, Chee Siang Leow, Junichi Hoshino, Takehito Utsuro, Hiromitsu Nishizaki</i>	
GGMDDC: An Audio Deepfake Detection Multilingual Dataset	1065
<i>Ravindra M. Purohit, Arth J. Shah, Hemant A. Patil</i>	
Robust Quantile Regression Under Unreliable Data	1071
<i>Yoshifumi Shoji, Masahiro Yukawa</i>	
Measurement of Relative Transfer Function for Own Voice in Head-Mounted Microphone Array.....	1077
<i>Kyoka Kazama, Taishi Nakashima, Nobutaka Ono</i>	
Dictionary Learning Based Two-Stage Near-Lossless Video Compression	1082
<i>Zuhai Zhang, Luheng Jia, Li Song, Shuyuan Zhu, Yuanfang Guo, Kebin Jia</i>	
EEND-EM: End-To-End Neural Speaker Diarization with EM-Network.....	1088
<i>Beom Jun Woo, Ji Won Yoon, Min Hyun Han, Chan Yeong Moon, Nam Soo Kim</i>	
Vocal Tract Length Perturbation-Based Pseudo-Speaker Augmentation Considering Speaker Variability for Speaker Verification	1093
<i>Hengyi Zou, Sayaka Shiota</i>	
Multi-Task Learning Approaches for Music Similarity Representation Learning Based on Individual Instrument Sounds.....	1099
<i>Takehiro Imamura, Yuka Hashizume, Tomoki Toda</i>	
Enhancing Early Plant Disease Detection: 1D to 2D Spectral Transformations	1105
<i>Mas Ira Syafila Mohd Hilmi Tan, Lai-Kuan Wong, Yuen Peng Loh, Chih-Yang Pee</i>	
EAViT: External Attention Vision Transformer for Audio Classification.....	1111
<i>Aquib Iqbal, Abid Hasan Zim, Md Asaduzzaman Tonmoy, Limengnan Zhou, Asad Malik, Minoru Kuribayashi</i>	
A Two-Stage Method for 3D Architecture Wireframe Reconstruction from Airborne LiDAR Point Cloud	1117
<i>Jiahao Zhang, Qi Liu, Le Hui, Yuchao Dai</i>	

Comparative Analysis of Voice Mimicry Attacks by High- And Low-Skilled Imitators on Speaker Verification Systems	1123
<i>Koji Iwano, Wakana Komuro, Manami Gomi</i>	
Performance Improvement of Single Plane-Wave Imaging Using U-Net and Discrete Wavelet Transform	1129
<i>Hiromi Shidara, Kanta Miura, Takuro Ishii, Koichi Ito, Takafumi Aoki, Yoshifumi Saijo, Jun Ohmiya</i>	
Compressed Deepfake Video Detection Based on 3D Spatiotemporal Trajectories.....	1135
<i>Zongmei Chen, Xin Liao, Xiaoshuai Wu, Yanxiang Chen</i>	
Ensemble Learning Based Head-Related Transfer Function Personalization Using Anthropometric Features	1143
<i>Yih-Liang Shen, Tai-Shi Chi</i>	
Secure Moving Object Detection Transformer in Compressed Video with Feature Fusion	1149
<i>Yuru Song, Yike Chen, Peijia Zheng, Yusong Du, Weiqi Luo</i>	
JAM: A Unified Neural Architecture for Joint Multi-Granularity Pronunciation Assessment and Phone-Level Mispronunciation Detection and Diagnosis Towards a Comprehensive CAPT System	1155
<i>Yue-Yang He, Bi-Cheng Yan, Tien-Hong Lo, Meng-Shin Lin, Yung-Chang Hsu, Berlin Chen</i>	
Optimizing Deep Q-Network for Shortest Path Computation of Mobile Robot Agents	1161
<i>A Sumarudin, Nana Sutisna, Infall Syafalni, Bambang Riyanto Trilaksono, Trio Adiono</i>	
Blind Self-Interference Analog Cancellor with Differential Delay for Backscatter Communications	1167
<i>Koichi Nishikawa, Shinsuke Ibi, Takumi Takahashi, Hisato Iwai</i>	
Efficient Adaptation for Real-World Omnidirectional Image Super-Resolution	1173
<i>Cuixin Yang, Rongkang Dong, Kin-Man Lam</i>	
Personal Voice Activity Detection with Ultra-Short Reference Speech	1179
<i>Longting Xu, Mingjun Zhang, Wenbin Zhang, Tianyi Wang, Jiawei Yin, Yu Gao</i>	
More Direct and Stage-Wise Network for Face Super Resolution	1185
<i>Yohei Horiguchi, Masaaki Ikehara, Kei Shibasaki</i>	
Data Augmentation Methods and Influence of Speech Recognition Performance for TED Talk's English to Japanese Speech Translation	1191
<i>Kento Masuda, Kazumasa Yamamoto, Seiichi Nakagawa</i>	
KhmerFormer: Multi-Scale CNNs-Transformer with External Attention for Ancient Khmer Palm Leaf Isolated Glyph Classification	1197
<i>Nimol Thuon, Jun Du</i>	
DDPMVC: Non-Parallel Any-To-Many Voice Conversion Using Diffusion Encoder	1203
<i>Ryuichi Hatakeyama, Kohei Okuda, Toru Nakashika</i>	
Training Deep Neural Networks with HSIC and Backpropagation.....	1209
<i>Roshan Birjais, Kevin I-Kai Wang, Waleed Abdulla</i>	
Real-Time Noise Estimation for Lombard-Effect Speech Synthesis in Human–Avatar Dialogue Systems.....	1214
<i>Yuto Ishikawa, Osamu Take, Tomohiko Nakamura, Norihiro Takamune, Yuki Saito, Shinnosuke Takamichi, Hiroshi Saruwatari</i>	

Zero-Shot Learning for Haze Removal Using Fusion of Near-Infrared and Color Images.....	1220
<i>Onhi Kato, Akira Kubota</i>	
A Study on Variable Embedding Locations of Reversible Spectral Speech Watermarking.....	1226
<i>Xuping Huang, Akinori Ito</i>	
Normalizing Flows-Based Latent Variable Rearrangement for Generative Image Steganography.....	1232
<i>Sifan Wu, Li Dong, Diqun Yan, Rangding Wang</i>	
Empower Typed Descriptions by Large Language Models for Speech Emotion Recognition.....	1238
<i>Haibin Wu, Huang-Cheng Chou, Kai-Wei Chang, Lucas Goncalves, Jiawei Du, Jyh-Shing Roger Jang, Chi-Chun Lee, Hung-Yi Lee</i>	
An Investigation on the Speech Recovery from EEG Signals Using Transformer.....	1244
<i>Tomoaki Mizuno, Takuya Kishida, Natsue Yoshimura, Toru Nakashika</i>	
WavLM and Omni-Scale CNNs: Enhancing Boundary Detection in Partially Spoofed Audio.....	1250
<i>Menghan Li, Zhihua Huang</i>	
NeRF-FCM: Feature Calibration Mechanisms for NeRF-Based 3D Object Detection.....	1255
<i>Hana Lebeta Goshu, Jun Xiao, Kin-Chung Chan, Cong Zhang, Mulugeta Tegegn Gameda, Kin-Man Lam</i>	
Complex CNN Incorporating Hilbert Transform for Steady-State Visual Evoked Potential BCI.....	1261
<i>Rintaro Takata, Yoshikazu Washizawa</i>	
MGVul: A Multi-Granularity Detection Framework for Software Vulnerability	1267
<i>Xiangyu Zhao, Yanjun Li, Zhengpeng Zha, Zhenhua Ling</i>	
Camera Focal Length Prediction for Neural Novel View Synthesis from Monocular Video	1273
<i>Dipanita Chakraborty, Werapon Chiracharit, Kosin Chamnongthai, Minoru Okada</i>	
Low-Resource Language Adaptation with Ensemble of PEFT Approaches.....	1278
<i>Chin Yuen Kwok, Sheng Li, Jia Qi Yip, Eng Siong Chng</i>	
Predictive Analysis of Driver Drowsiness Progression: Multi-Level Drowsiness Classification Using Physiological Signals	1284
<i>Natchira Dachoponchai, Yodchanan Wongsawat, Jetsada Arnin</i>	
Color Enhancement for the Colorblind Using Color Correction Intensity Map and Pix2pix Image Conversion.....	1290
<i>Shu Komatsu, Akira Kubota</i>	
ACE-Flow: Auto Color Encoding for Enhanced Low-Light Image Restoration.....	1295
<i>Jiachen Qiu, Yushen Zuo, Kin-Man Lam</i>	
Electroencephalogram-Based Effective Features for Sustained Attention Assessment in Conversation.....	1301
<i>Masaya Togashi, Ingon Chanpornpakdi, Toshihisa Tanaka</i>	
Leveraging Attention Mechanisms for Breast Cancer Diagnosis	1307
<i>Akumalla Brahma Reddy, Bach-Tung Pham, Tung-Yu Zhuang, Bima Paristao, Pao-Chi Chang, Jia-Ching Wang</i>	
Leveraging IoT and Machine Learning for Efficient Rice Stock Monitoring and Prediction	1311
<i>Nana Sutisna, Aditya Prawira Nugroho, Christopher Jeffrey, Patrick Amadeus Irawan, Rizky Ramadhana, Ronggur Mahendra, Michael Jonathan, Infall Syafalni, Trio Adiono</i>	

Blind Estimation of Room Volume from Reverberant Speech Based on the Modulation Transfer Function.....	1317
<i>Nutchanon Siripool, Suradej Duangpummet, Jessada Karnjana, Waree Kongprawechanon, Masashi Unoki</i>	
Semi-Supervised Far-Field Speaker Verification with Distance Metric Domain Adaptation	1323
<i>Han Wang, Mingrui He, Mingjun Zhang, Longting Xu</i>	
Synchronization of Signals with Sampling Rate Offset and Missing Data Using Dynamic Programming Matching.....	1329
<i>Hayato Takeuchi, Takao Kawamura, Nobutaka Ono, Shoko Araki</i>	
IoT-Based Smart Attendance System Using Face Recognition and Motion Detection	1335
<i>Umi Syamimi, Lim Chern Hong, Lillian Yee Kiaw Wang</i>	
Robust Adaptive Filtering Based on Adaptive Projected Subgradient Method: Moreau Enhancement of Distance Function.....	1341
<i>Daiki Sawada, Masahiro Yukawa</i>	
Non-Target Conversion Based Speech Steganography for Secure Speech Communication System	1347
<i>Mingjun Zhang, Yan Feng, Yu Gao, Longting Xu</i>	
A Multi-Domain Camera Model Identification Feature Restoration Network to Counter AI Compression Attacks	1353
<i>Jinkai Zhang, Zijuan Han, Yunxia Liu, Yang Yang</i>	
Deep Learning-Based Intraoperative Video Analysis for Cataract Surgery Instrument Identification	1359
<i>Z. Guo, Y. H. Chan, N. F. Law</i>	
Implementation of a Robot Operation System-Based Network for Sound Source Localization Using Multiple Drones.....	1366
<i>Takumi Yamamoto, Kotaro Hoshiba, Benjamin Yen, Kazuhiro Nakadai</i>	
Disentangling Speaker Representations from Intuitive Prosodic Features for Speaker-Adaptive and Prosody-Controllable Speech Synthesis	1372
<i>Pengyu Cheng, Zhenhua Ling, Meng Meng, Yujun Wang</i>	
A Joint Graph Signal and Laplacian Denoising Network.....	1378
<i>Zepeng Zhang, Ziping Zhao</i>	
Annotation-Free Fine-Tuning for Unsupervised Anomalous Sound Detection.....	1383
<i>Kai Guo, Xiang Xie, Fengrun Zhang</i>	
Comparative Analysis of Glottal and Vocal Tract Features in Dysarthria	1388
<i>Geeta Sai Sahasra, Kadwasra Swapna, Arushi Srivastava, Aditya Pusuluri, Hemant A. Patil</i>	
Contrast-Aware DCT for Image Enhancement with JPEG Compatible Coding.....	1394
<i>Kohei Hayashi, Soichiro Honda, Hirokazu Kamei, Yoshihiro Maeda, Norishige Fukushima</i>	
Comparative Evaluation of Fine-Tuned Hybrid Transformer and Band-Split Recurrent Neural Networks for Music Source Separation.....	1400
<i>Ken Kalang Al Qalyubi, Nur Ahmadi, Dessi Puji Lestari</i>	
Dysarthria Severity Classification Using Phase Based Features of LP Residual.....	1405
<i>Rohini Sri Mannepalli, Aditya Pusuluri, Hemant A. Patil</i>	

Significance of Entropy Based Features for Dysarthric Severity Level Classification	1410
<i>Meghana Avula, Aditya Pusuluri, Hemant A. Patil</i>	
Feature Extraction for Machine Learning-Based Sleep Stage Classification Using PPG-Derived Parameters and Skin Temperature	1416
<i>Suphachok Buaruk, Sikawat Thanaviratananich, Peerasit Treesuthacheep, Somrudee Deepaisarn</i>	
Enhancing Acoustic Scene Classification with Layer-Wise Fine-Tuning on the SSAST Model	1422
<i>Shuting Hao, Daisuke Saito, Nobuaki Minematsu</i>	
Non-Blind Deblurring Using Probabilistic Models and Spatial Adaptive Restoration.....	1428
<i>Chun-Lin Liao, Jian-Jiun Ding, Chun-Jen Shih</i>	
A Pilot Study of Applying Sequence-To-Sequence Voice Conversion to Evaluate the Intelligibility of L2 Speech Using a Native Speaker's Shadowings	1434
<i>Haopeng Geng, Daisuke Saito, Nobuaki Minematsu</i>	
EADSum: Element-Aware Distillation for Enhancing Low-Resource Abstractive Summarization	1440
<i>Jia-Liang Lu, Bi-Cheng Yan, Yi-Cheng Wang, Tien-Hong Lo, Hsin-Wei Wang, Li-Ting Pai, Berlin Chen</i>	
Enhanced Detection of Illegally Parked Vehicles Using YOLO and Good Feature to Track Methods	1446
<i>Fauzan Maftuh Alwafi, Bobby Mugi Pratama, Phuong Thi Le, Bima Prihasto, Jia-Ching Wang</i>	
LEAD Dataset: How Can Labels for Sound Event Detection Vary Depending on Annotators?	1452
<i>Naoki Koga, Yoshiaki Bando, Keisuke Imoto</i>	
A Tiny Whisper-SER: Unifying Automatic Speech Recognition and Multi-Label Speech Emotion Recognition Tasks.....	1458
<i>Huang-Cheng Chou</i>	
Experimental Evaluation of Speech Enhancement for In-Car Environment Using Blind Source Separation and DNN-Based Noise Suppression.....	1464
<i>Yutsuki Takeuchi, Taishi Nakashima, Nobutaka Ono, Takashi Takazawa, Shuhei Shimanoe, Yoshinori Tsuchiya</i>	
Multi-Observed Authentication: A Secure and Usable Authentication Based on Multi-Point Observation of a Single Physical Credential	1470
<i>Wataru Hatakeyama, Shinnosuke Nozaki, Ayumi Serizawa, Mizuho Yoshihira, Masahiro Fujita, Ayako Yoshimura, Tetsushi Ohki, Masakatsu Nishigaki</i>	
Automated Pseudo-Label Generation and Parallel Computing for Enhanced Few-Shot Medical Image Segmentation	1476
<i>Trong-Duc Nguyen, Tien-Dung Do, Thanh-Ha Do</i>	
PBJDT: Point-Based Joint Detection-And-Tracking	1482
<i>Zhen-Xun Lee, Jian-Jiun Ding</i>	
Drone Audition: Dataset and Methods for Ground Surface Material Classification Using Drone Noise in Outdoor Environment.....	1488
<i>Tsubasa Yano, Benjamin Yen, Kazuhiro Nakadai</i>	
Auxiliary-Function-Based Steering Vector Estimation Method for Spatially Regularized Independent Low-Rank Matrix Analysis.....	1494
<i>Sota Hirata, Norihiro Takamune, Kouei Yamaoka, Daichi Kitamura, Hiroshi Saruwatari, Yu Takahashi, Kazunobu Kondo</i>	

Multi-Band Satellite Image Analysis for Multi-Label Classification	1500
<i>Sarah Shamina Abdul Rauf, Mas Ira Syafila Mohd Hilmi Tan, Yuen Peng Loh</i>	
Incorporating Auditory Processing into Undergraduate Signal Processing Courses to Enhance Student Learning	1506
<i>Kaibao Nie</i>	
A Comparative Study on the Biases of Age, Gender, Dialects, and L2 Speakers of Automatic Speech Recognition for Korean Language	1511
<i>Jonghwan Na, Yeseul Park, Bowon Lee</i>	
Detecting Spoof Voices in Asian Non-Native Speech: An Indonesian and Thai Case Study	1517
<i>Aulia Adila, Candy Olivia Mawalim, Masashi Unoki</i>	
Two-Stage Framework for Robust Speech Emotion Recognition Using Target Speaker Extraction in Human Speech Noise Conditions	1523
<i>Jinyi Mi, Xiaohan Shi, Ding Ma, Jiajun He, Takuya Fujimura, Tomoki Toda</i>	
Enhanced Sparse Convolutional Detection Model for 3D Object Detection in Autonomous Vehicles Adapted to Traffic Conditions in Vietnam.....	1529
<i>Vu Hoang Dung, Nguyen Trung Kien, Do Thanh Ha</i>	
SSL-Based Chewing and Swallowing Detection Using Multiple Skin-Contact Microphones.....	1535
<i>Toshihiro Tsukagoshi, Kazuhiro Koizumi, Masafumi Nishida, Masafumi Nishimura</i>	
Data Generation for Speaker Diarization by Speaker Transition Information.....	1540
<i>Keigo Ichikawa, Sei Ueno, Akinobu Lee</i>	
LoFLAT: Local Feature Matching Using Focused Linear Attention Transformer	1545
<i>Naijian Cao, Renjie He, Yuchao Dai, Mingyi He</i>	
Capturing Dynamic Identity Features for Speaker-Adaptive Visual Speech Recognition.....	1551
<i>Sara Kashiwagi, Keitaro Tanaka, Shigeo Morishima</i>	
Cross Lingual Speech Representation for Infant Cry Classification	1557
<i>Hiya Chaudhari, Arth J. Shah, Hemant A. Patil</i>	
Enhancing Cell Segmentation Using Deep Learning Models by Custom Processing Techniques	1562
<i>Van-De Nguyen, Minh-Huong Hoang Dang, Quang-Huy Nguyen, Manh-Cuong Dinh, Thanh-Ha Do</i>	
Seismic-Ionospheric Precursor Prediction Using Deep Learning.....	1567
<i>Bach-Tung Pham, Pao-Chi Chang, Jia-Ching Wang</i>	
Pop Noise Detection Using Group Delay Cepstral Coefficients	1571
<i>Arth Shah, Prathav Kevadiya, Hemant A. Patil</i>	
Inference Efficient Source Separation Using Input-Dependent Convolutions	1577
<i>Shogo Seki, Li Li</i>	
NecoBERT: Self-Supervised Learning Model Trained by Masked Language Modeling on Rich Acoustic Features Derived from Neural Audio Codec	1582
<i>Wataru Nakata, Takaaki Saeki, Yuki Saito, Shinnosuke Takamichi, Hiroshi Saruwatari</i>	
Privacy-Preserving Anomaly Detection in Bitstream Video Based on Gaussian Mixture Model	1588
<i>Yike Chen, Yuru Song, Peijia Zheng, Yusong Du, Weiqi Luo</i>	

Targeted Representation with Information Disentanglement Encoding Networks.....	1594
<i>Takumi Nagawaki, Keisuke Ikeda, Kohei Chike, Hiroyuki Nagano, Masaki Nose, Satoshi Tamura</i>	
High and Low Frequency Region Separation Method for Adaptive Image Expansion.....	1599
<i>Shao-Yun Luo, Kuei-Chen Chen, Jian-Jiun Ding, Cheng-Che Lee, Hsin-Jung Lee</i>	
Unleashing Attributes-Content Adaptation with Multi-Color Spaces for Food Photo Aesthetic Assessment.....	1605
<i>Shintami Chusnul Hidayati, Muhammad Valda Rizky Nur Firdaus, Riki Wahyu Nur Dianto, Sarwosri</i>	
PG-MDD: Prompt-Guided Mispronunciation Detection and Diagnosis Leveraging Articulatory Features	1611
<i>Meng-Shin Lin, Bi-Cheng Yan, Tien-Hong Lo, Hsin-Wei Wang, Yue-Yang He, Wei-Cheng Chao, Berlin Chen</i>	
An Explainable Raman Spectral Classification Pipeline Via NMF and SHAP: A Case Study of Pen Ink Colors.....	1617
<i>Pongpon Lapsatid, Somrudee Deepaisarn, Pitak Eiamchai</i>	
GSBIQA: Green Saliency-Guided Blind Image Quality Assessment Method	1623
<i>Zhanxuan Mei, Yun-Cheng Wang, C.-C. Jay Kuo</i>	
Scene-Segmentation-Based Exposure Compensation for Tone Mapping of High Dynamic Range Scenes.....	1629
<i>Yuma Kinoshita, Hitoshi Kiya</i>	
EMO-Codec: An In-Depth Look at Emotion Preservation Capacity of Legacy and Neural Codec Models with Subjective and Objective Evaluations	1635
<i>Wenze Ren, Yi-Cheng Lin, Huang-Cheng Chou, Haibin Wu, Yi-Chiao Wu, Chi-Chun Lee, Hung-Yi Lee, Hsin-Min Wang, Yu Tsao</i>	
A Real-Time Platform for Portable and Scalable Active Noise Mitigation for Construction Machinery.....	1641
<i>Woon-Seng Gan, Santi Peksi, Chung Kwan Lai, Yen Theng Lee, Dongyuan Shi, Bhan Lam</i>	
Pressure Matching Using Data-Driven Estimation for Sound Fields and Transfer Functions.....	1647
<i>Koki Horikoshi, Gen Sato, Izumi Tsunokuni, Yusuke Ikeda</i>	
Analytic Study of Text-Free Speech Synthesis for Raw Audio Using a Self-Supervised Learning Model	1652
<i>Joonyong Park, Daisuke Saito, Nobuaki Minematsu</i>	
Acoustic Model Adaptation in Noisy and Reverberated Scenarios Using Multi-Task Learned Embeddings	1658
<i>Aditya Raikar, Meet Soni, Ashish Panda, Sunil Kumar Kopparapu</i>	
Generating Room Impulse Responses Using Neural Networks Trained with Weighted Combinations of Acoustic Parameter Loss Functions	1663
<i>Hualin Ren, Christian Ritz, Jiahong Zhao, Xiguang Zheng, Daeyoung Jang</i>	
GVDIE: A Zero-Shot Generative Information Extraction Method for Visual Documents Based on Large Language Models	1669
<i>Siyang Qi, Fei Wang, Hongzhi Sun, Yang Ge, Bo Xiao</i>	

Investigating the Language Independence of Voice Activity Projection Models Through Standardization of Speech Segmentation Labels	1675
<i>Yuki Sato, Yuya Chiba, Ryuichiro Higashinaka</i>	
Confidence-Aware Learning for Person Re-Identification with Noisy Labels	1681
<i>Duhyun Kim, Jae-Young Sim</i>	
Generalized SpecAugment: Robust Online Augmentation Technique for End-To-End Automatic Speech Recognition	1686
<i>Meet Soni, Ashish Panda, Sunil Kumar Kopparapu</i>	
A Byte-Based GPT-2 Model for Bit-Flip JPEG Bitstream Restoration	1691
<i>Hao Qin, Haoran Sun, Yi Wang</i>	
A Preliminary Study on Analysing Mandarin Tone Values of Romance L2 Mandarin Learners	1697
<i>Wu-Hao Li, Te-Hsin Liu, Chen-Yu Chiang</i>	
Enhancing Remote Adversarial Patch Attacks on Face Detectors with Tiling and Scaling	1703
<i>Masora Okano, Koichi Ito, Masakatsu Nishigaki, Tetsushi Ohki</i>	
GMNER-LF: Generative Multi-Modal Named Entity Recognition Based on LLM with Information Fusion	1709
<i>Huiyun Hu, Junda Kong, Fei Wang, Hongzhi Sun, Yang Ge, Bo Xiao</i>	
Novel Estimators for the Number of Susceptible Individuals in SIR Models of Infectious Epidemics	1715
<i>Michaël Antonie Van Wyk, André Martin McDonald, David M. Rubin, Fangfang Zhang</i>	
Swarm Active Audition System with Robots and Drones for a Search and Rescue Task.....	1721
<i>Kazuhiro Nakadai, Makoto Kumon, Yoko Sasaki, Kotaro Hoshiba, Benjamin Yen</i>	
PPHiFi-TTS: Phonetic Preserved High-Fidelity Text-To-Speech for Long-Term Speech Dependencies.....	1727
<i>Ravindrakumar M. Purohit, Dharmendra H. Vaghera, Arth J. Shah, Hemant A. Patil</i>	
Test-Time Optimization for Post-Processing of Compressed Videos	1733
<i>Hongil Kim, Changwoo Han, Donghyun Kim, Sung-Chang Lim, Seung-Won Jung</i>	
Peer Learning Via Shared Speech Representation Prediction for Target Speech Separation.....	1739
<i>Xusheng Yang, Zifeng Zhao, Yuexian Zou</i>	
Physics-Informed Neural Networks for Estimation of Scattered Sound Fields with Boundary Condition.....	1746
<i>Ryosuke Onizawa, Gen Sato, Izumi Tsunokuni, Yusuke Ikeda</i>	
WildPose: HRNet-Based Lightweight and Efficient Wildlife Pose Estimation.....	1751
<i>Sibusiso Reuben Bakana, Yongfei Zhang, Bhekisipho Twala</i>	
Teager Energy Cepstral Coefficients for Audio Deepfake Detection	1757
<i>Ritik Mahyavanshi, C. V. Mahesh Reddy, Arth J. Shah, Hemant A. Patil</i>	
META: Text Detoxification by Leveraging METAmorphic Relations and Deep Learning Methods.....	1763
<i>Alika Choo, Arghya Pal, Sailaja Rajanala, Arkendu Sen</i>	
Parameterizing Hierarchical Particle Filters with Concept Drift for Time-Varying Parameter Estimation.....	1769
<i>Joshua Murphy, Conor Rosato, Andrew Millard, Simon Maskell</i>	

Differences Between Singer and Speaker Verification: Training Singer Feature Representation Extractor Utilizing Singing Voice Characteristics	1775
<i>Sayaka Toma, Tomoki Ariga, Yosuke Higuchi, Ichiju Hayasaka, Rie Shigyo, Tetsuji Ogawa</i>	
Developing a Multilingual Spontaneous L2 Speech Corpus for Automated Proficiency Assessment	1780
<i>Seunghye Han, Sunhee Kim, Minhwa Chung</i>	
Data-Driven Physics-Informed Neural Network for Sound Field Estimation in Rooms of Arbitrary Size	1786
<i>Gen Sato, Yusuke Ikeda</i>	
Lifelong Person Re-Identification with Backward-Compatibility	1791
<i>Minyoung Oh, Jae-Young Sim</i>	
Towards a B-Format Ambisonic Room Impulse Response Generator Using Conditional Generative Adversarial Network	1797
<i>Hualin Ren, Christian Ritz, Jiahong Zhao, Xiguang Zheng, Daeyoung Jang</i>	
GPGAN-VC: Enhancing Voice Conversion Using Gradient Penalty	1803
<i>Ravindrakumar M. Purohit, Dharmendra H. Vaghera, Hemant A. Patil</i>	
Improved Cassava Plant Disease Classification with Leaf Detection	1809
<i>Ming Xuan Chai, Yao Deng Fam, Quinito Norman Octaviano, Chih-Yang Pee, Lai-Kuan Wong, Mas Ira Syafila Mohd Hilmi Tan, John See</i>	
A Study on Packet-Level Index Modulation Using Frequency Offsets Within a LoRaWAN Channel	1815
<i>Mai Ohta, Hiroki Matsuura, Takeo Fujii</i>	
Audio Similarity Detection.....	1821
<i>Siddharth Harsh Y. Malhotra, Sapan H. Mankad</i>	
A Document Presentation Attack Detection Scheme with Optical Flow Under a Flashlight	1827
<i>Changsheng Chen, Wenyu Chen, Ximin Chen, Haodong Li</i>	
Development and Evaluation of a Semi-Autonomous Parallel Attentive Listening System	1833
<i>Divesh Lala, Koji Inoue, Haruki Kawai, Zi Haur Pang, Mikey Elmers, Tatsuya Kawahara</i>	
Prediction of Negative User Reactions Towards System Responses During Attentive Listening	1839
<i>Divesh Lala, Koji Inoue, Tatsuya Kawahara</i>	
Data Selection Using Spoken Language Identification for Low-Resource and Zero-Resource Speech Recognition	1845
<i>Jianan Chen, Chenhui Chu, Sheng Li, Tatsuya Kawahara</i>	
What to Refer and How? - Exploring Handling of Auxiliary Information in Target Speaker Extraction	1851
<i>Tomohiro Hayashi, Riku Ogino, Kohei Saijo, Tetsuji Ogawa</i>	
New Approach on Smiling Faces with Domain Transfer in Latent Space.....	1857
<i>Felix Ming-Fei Duan, Wan-Chi Siu, Chun-Chuen Hui</i>	
Robust Image Watermarking Scheme Under Halftone Distortion with Surrogate Model	1862
<i>Changsheng Chen, Xijin Li</i>	
High-Quality Facial Pose Generation with Latent Space Processing	1868
<i>Wing-Ho Cheng, Wan-Chi Siu, H. Anthony Chan</i>	

Visual Semantic Alignment Network Based on Pre-Trained ViT for Few-Shot Image Classification	1875
<i>Jiaming Zhang, Jijie Wu, Xiaoxu Li</i>	
Agent Attention Feature Reconstruction Network for Fine-Grained Few-Shot Image Classification.....	1881
<i>Dongfei Chang, Jijie Wu, Xiaoxu Li</i>	
Adaptive Spatial Re-Sampling Method for Video Coding for Machines	1887
<i>Eun-Bin An, Ayoung Kim, Soon-Heung Jung, Hyon-Gon Choo, Kwang-Deok Seo</i>	
A Multi-Perceptual Learning Network for Retina OCT Image Denoising and Classification	1891
<i>Zhe Xiao, Zongqi He, Zhuoning Xu, Yunze Li, Zelin Song, Calvin Leighton, Li Wang, Shanru Liu, Shiun Yee Wong, Wenfeng Huang, Wenjing Jia, Kin-Man Lam</i>	
SRC-GAudio: Sampling-Rate-Controlled Audio Generation.....	1897
<i>Chenxing Li, Manjie Xu, Dong Yu</i>	
Source Attribution for Images Generated by Diffusion-Based Text-To-Image Models: Exploring the Forensics Approach	1903
<i>Xinqi Jiang, Jinyu Tian</i>	
On the Importance of Time and Pitch Relativity for Transformer-Based Symbolic Music Generation	1909
<i>Tatsuro Inaba, Kazuyoshi Yoshii, Eita Nakamura</i>	
Multi-Modal Video Summarization Based on Two-Stage Fusion of Audio, Visual, and Recognized Text Information	1915
<i>Zekun Yang, Jiajun He, Tomoki Toda</i>	
AFSDet: Video Small Object Detection Based on Adaptive Focused Slicing.....	1921
<i>Kangjian Huang, Yan Yang, Yongquan Jiang, Xiaobo Zhang, Zhuyi Angelina Li</i>	
Diverse Time-Frequency Attention Neural Network for Acoustic Echo Cancellation	1927
<i>Jinzhao Yao, Hongqing Liu, Yi Zhou, Lu Gan, Junkang Yang</i>	
LDMSE: Low Computational Cost Generative Diffusion Model for Speech Enhancement.....	1933
<i>Yuki Nishi, Koichi Shinoda, Koji Iwano</i>	
Continual Learning with Self-Organizing Maps: A Novel Group-Based Unsupervised Sequential Training Approach.....	1939
<i>Gaurav Hirani, Kevin I-Kai Wang, Waleed Abdulla</i>	
Improving Semi-Supervised Object Detection by ROI-Enhanced Contrastive Learning.....	1945
<i>Teng-Kuan Huang, Mei-Chen Yeh</i>	
Dual Motion Attention and Enhanced Knowledge Distillation for Video Frame Interpolation.....	1951
<i>Dengyong Zhang, Runqi Lou, Jiaxin Chen, Xin Liao, Gaobo Yang, Xiangling Ding</i>	
Detecting Coronary Artery Stenosis from Cardiac CT Images Using 3D CNNs	1957
<i>Masaki Aono, Tetsuya Asakawa, Kazuki Shimizu, Masashi Hahsimoto, Takeshi Miyaji, Kei Nomura</i>	
Real-Time Segmentation of Coronary Artery Calcification Using Spatial Attention and Parallel Convolution	1963
<i>Tetsuya Asakawa, Masashi Hashimoto, Takeshi Miyaji, Kazuki Shimizu, Kei Nomura, Masaki Aono</i>	

Effective Speech Data Augmentation Method to Improve Customer Service Representative Speech Recognition System Performance	1968
<i>Huiyong Bak, Changhyeon Jeong</i>	
LLM as Decoder: Investigating Lattice-Based Speech Recognition Hypotheses Rescoring Using LLM	1973
<i>Sheng Li, Yuka Ko, Akinori Ito</i>	
EavaNet: Enhancing Emotional Facial Expressions in 3D Avatars Through Speech-Driven Animation.....	1978
<i>Seyun Um, Yongju Lee, Wooseok Ko, Yuan Zhou, Sangyoun Lee, Hong-Goo Kang</i>	
Joseph: Phonetic-Aware Speaker Embedding for Far-Field Speaker Verification.....	1984
<i>Zezhong Jin, Youzhi Tu, Man-Wai Mak</i>	
Accelerated Real-Time Local Maxima Detection in Video Streams Using FPGA Technology	1990
<i>Anindhita Nayazirly Sukarno, Yahwista Salomo, Trio Adiono, Infall Syafalni, Nana Sutisna, Rahmat Mulyawan</i>	
Context-FFT: A Context Feed Forward Transformer Network for EEG-Based Speech Envelope Decoding	1996
<i>Ximin Chen, Yuting Ding, Nan Yan, Changsheng Chen, Fei Chen</i>	
Optimal Investment with Incomplete Information and Herd Effect	2001
<i>Huisheng Wang, Mingxiao Liu, Ji Qi, H. Vicky Zhao</i>	
A Two-Stage Wall Parameters Estimation Algorithm for MIMO Through-The-Wall Radar.....	2007
<i>Zhirun Li, Shisheng Guo, Jiahui Chen, Zhihao Zhu, Chenqiu, Guolong Cui, Yutao Xiang</i>	
Bluemarble: Bridging Latent Uncertainty in Articulatory-To-Speech Synthesis with a Learned Codebook	2012
<i>Seyun Um, Miseul Kim, Doyeon Kim, Hong-Goo Kang</i>	
MTFNet: Multi-Scale Transformer Framework for Robust Emotion Monitoring in Group Learning Settings.....	2018
<i>Yi Zhang, Fangyuan Liu, Jiajia Song, Qi Zeng, Hui He</i>	
Target Speaker Extraction Method by Emphasizing the Active Speech with an Additional Enhancer	2026
<i>Xue Yang, Changchun Bao, Xu Zhang, Xianhong Chen</i>	
GE2E-AC: Generalized End-To-End Loss Training for Accent Classification	2032
<i>Chihiro Watanabe, Hirokazu Kameoka</i>	
Efficient Feature Selection for Word Embedding Dimension Reduction	2038
<i>Jintang Xue, Yun-Cheng Wang, Chengwei Wei, C.-C. Jay Kuo</i>	
Tiny Object Detection Enhancement for Large-Scale Remote Sensing Imagery	2044
<i>Tianwei Zhang, Lianru Gao, Xu Sun, Lina Zhuang</i>	
Scale-Invariant Online Voice Activity Detection Under Various Environments	2049
<i>Ryu Takeda, Kazunori Komatani</i>	
Fine-Grained Quantitative Emotion Editing for Speech Generation	2055
<i>Sho Inoue, Kun Zhou, Shuai Wang, Haizhou Li</i>	

YOLO-DC: Enhancing Object Detection with Deformable Convolutions and Contextual Mechanism	2061
<i>Dengyong Zhang, Chuanzhen Xu, Jiaxin Chen, Bin Deng, Xin Liao</i>	
One-Step Spectral Estimation for Euclidean Distance Matrix Approximation	2067
<i>Yicheng Li, Xinghua Sun</i>	
Handling Missing Data in Limited-View Photoacoustic Tomography Using Compressive Sensing Algorithm-Based Deep Learning.....	2073
<i>Mary John, Imad Barhumi</i>	
Green Video Camouflaged Object Detection	2079
<i>Xinyu Wang, Hong-Shuo Chen, Zhiruo Zhou, Suyu You, Azad M. Madni, C.-C. Jay Kuo</i>	
Forward Prediction-Guided Cross-Partition Targeted Pruning for VVenC.....	2085
<i>Jingyuan Tang, Songlin Sun</i>	
Marker-Aware Ovarian Tumor Segmentation from Ultrasound Images.....	2091
<i>Hoang-Son Bui, Sy-Hoang Tran, Thuy-Binh Nguyen, Thanh-Hai Tran, Hai Vu, Thi-Lan Le</i>	
Enhancing Semiconductor X-RAY Images: A Framework Combining Denoising and Super-Resolution Modules with a Novel Dataset	2097
<i>Jae Hoon Shim, Min Woo Kim, Tae Gyu Lim, Byung Seok Min, Sang Hwa Lee, Nam Ik Cho</i>	
Improving Speaker Consistency in Speech-To-Speech Translation Using Speaker Retention Unit-To-Mel Techniques	2103
<i>Rui Zhou, Akinori Ito, Takashi Nose</i>	
Keyword Spotting for Dialectal Speech and Introduction of Wav2vec2.0	2109
<i>Tomohiro Ariga, Reo Minakawa, Kazunori Kojima, Shi-Wook Lee, Yoshiaki Itoh</i>	
LCMV-Based Scan-And-Sum Beamforming for Region Source Extraction.....	2114
<i>Aoto Yasue, Benjamin Yen, Katsutoshi Itoyama, Kazuhiro Nakadai</i>	

Author Index