

Full Program

Session		Room	Chair
Overview Session 1		Meeting Room 1	
Date	Time	Title	Speaker
4-Dec	16:20-16:40	A Decade of Progress in Sound Event Localization and Detection: Transforming Environmental Sound Analysis for Real-World Impact	Woon-Seng Gan, Nanyang Technological University
	16:40-17:00	Exploring the Forward-Forward Algorithm: A Novel Learning Approach	Waleed H. Abdulla, The University of Auckland
	17:00-17:20	Eye-gaze-based Human-Intention Detection	Kosin Chamnongthai, King Mongkut's University of Technology Thonburi
	17:20-17:40	From GPT Evolution to Enterprise Deployment: Key Trends in Generative AI	Jing-Ming Guo, National Taiwan University of Science and Technology
	17:40-18:00	An Overview of Online Distributed Kernel Methods for Supervised and Unsupervised Learning	Anthony Kuh, University of Hawaii

Session		Room	Chair
Overview Session 2		Meeting Room 8	
Date	Time	Title	Speaker
5-Dec	10:20-10:40	An AI-based Diagnostic-aid for Epileptic Electroencephalography	Toshihisa Tanaka, Tokyo University of Agriculture and Technology

	10:40-11:00	Machine Learning for Analytics Architecture: AI to Design AI Video	Chris Gwo Giun Lee, National Cheng Kung University
	11:00-11:20	Compression of Large AI Models	Weisi Lin, Nanyang Technological University
	11:20-11:40	Introduction to Multi-Camera Systems and 3D Quality Assessment	Sanghoon Lee, Yonsei University
	11:40-12:00	Highlight of New Image Generative Models and Applications to Image Manipulations	Wan-Chi Siu, Hong Kong Polytechnic University & St. Francis University

Session		Room	Chair
Overview Session 3		Merged Room (Room 10 + 11)	
Date	Time	Title	Speaker
6-Dec	9:00-9:20	Overview of Source Camera Identification Techniques	Bonnie N. F. Law, The Hong Kong Polytechnic University
	9:20-9:40	Recent Advances in Complete Quality Preserving Data Hiding	KokSheik Wong, Monash University Malaysia
	9:40-10:00	Real or Fake? Frontiers of Countering Fake Media in the Age of Infodemics	Isao Echizen, National Institute of Informatics
	10:00-10:20	User Preference Modeling and Analysis in Choice Problems	H. Vicky Zhao, Tsinghua University

Oral Sessions

Session	Room	Chair
---------	------	-------

Machine Learning and Data Analytics		Meeting Room 1	-
Date	Time	Title	Authors
04-12-2024	11:00-11:20	Psychological Driving Style Estimation from GPS Sensor Data Alone	Horimoto, Hiroto; Kimura, Ryusei; Tanaka, Takahiro; Okada, Shogo*
	11:20-11:40	Adversarial Augmentation and Adaptation for Speech Recognition	Chien, Jen-Tzung*; Sun, Wei-Yu
	11:40-12:00	Empathetic Response Generation via Regularized Q-Learning	Chien, Jen-Tzung*; Wu, Yi-Chien
	12:00-12:20	Continual Learning with Self-Organizing Maps: A Novel Group-Based Unsupervised Sequential Training Approach	Hirani, Gaurav R*; Wang, Kevin I-Kai; Abdulla, Waleed

Session		Room	Chair
Machine Learning and Data Analytics		Meeting Room 2	-
Date	Time	Title	Authors
04-12-2024	11:00-11:20	YOLO for High Resolution Images without Retraining	Minami, Daisuke*; Nishikawa, Kiyoshi
	11:20-11:40	Noise-Robust Estimation of Early-part Room Impulse Responses based on Physics-Informed Neural Network with Dynamic Pulling Method	Kurata, Ken*; Sato, Gen; Tsunokuni, Izumi; Ikeda, Yusuke
	11:40-12:00	A Multi-Domain Camera Model Identification Feature Restoration Network to Counter AI Compression Attacks	jinkai, zhang*

	12:00-12:20	Deep Learning-based Intraoperative Video Analysis for Cataract Surgery Instrument Identification	Guo, Zhe*; Chan, Yuk Hee; Law, Ngai Fong
--	-------------	--	--

Session		Room	Chair
Image, Video, and Multimedia		Meeting Room 3	-
Date	Time	Title	Authors
04-12-2024	11:00-11:20	GSBIQA: Green Saliency-guided Blind Image Quality Assessment Method	Mei, Zhanxuan*; Wang, Yun-Cheng; Kuo, C.-C. Jay
	11:20-11:40	AFSDet: Video Small Object Detection Based on Adaptive Focused Slicing	Huang, Kangjian; Yang, Yan*; Jiang, Yongquan; Zhang, Xiaobo; Li, Zhuyi Angelina
	11:40-12:00	Dual Motion Attention and Enhanced Knowledge Distillation for Video Frame Interpolation	Zhang, Deng yong*; lou, runqi; Chen, Jiaxin; Liao, Xin; Yang, Gaobo; ding, xiangling
	12:00-12:20	EavaNet: Enhancing Emotional Facial Expressions in 3D Avatars through Speech-Driven Animation	um, seyun*; Lee, YongJu; Ko, WooSeok; Zhou, Yuan; Lee, Sangyoun; Kang, Hong-Goo

Session		Room	Chair
Signal and Information Processing & Systems		Meeting Room 4	-
Date	Time	Title	Authors
04-12-2024	11:00-11:20	On the Importance of Time and Pitch Relativity for	Inaba, Tatsuro*; Yoshii, Kazuyoshi; Nakamura, Eita

		Transformer-based Symbolic Music Generation	
	11:20-11:40	Optimal Investment With Incomplete Information and Herd Effect	Wang, Huisheng; Liu, Mingxiao; Qi, Ji; Zhao, H. Vicky*
	11:40-12:00	YOLO-DC: Enhancing object detection with deformable convolutions and contextual mechanism	Zhang, Deng yong*; Xu, Chuazhen; Chen, Jiabin; Liao, Xin
	12:00-12:20	One-step Spectral Estimation for Euclidean Distance Matrix Approximation	Li, Yicheng*; Sun, Xinghua

Session		Room	Chair
Speech and Language Processing		Meeting Room 5	-
Date	Time	Title	Authors
04-12-2024	11:00-11:20	SDNet: Noise-Robust Bandwidth Extension under Flexible Sampling Rates	Yang, Junkang*; Liu, Hongqing; Gan, Lu; Zhou, Yi; Li, Xing; Jia, Jie; Yao, Jinzhuo
	11:20-11:40	GLASS: Investigating Global and Local context Awareness in Speech Separation	Ho, Kuan-Hsun*; Yu, En-Lun; Hung, Jieh-wei; Huang, Shih-Chieh; Chen, Berlin
	11:40-12:00	Low-resource Language Adaptation with Ensemble of PEFT Approaches	Kwok, Chin Yuen*; Li, Sheng; Yip, Jia Qi; Chng, Eng Siong
	12:00-12:20	Diverse Time-Frequency Attention Neural Network for Acoustic Echo Cancellation	Yao, Jinzhuo*; Liu, Hongqing; Zhou, Yi; Gan, Lu; Yang, Junkang

Session	Room	Chair
---------	------	-------

Audio Processing		Meeting Room 6	-
Date	Time	Title	Authors
04-12-2024	11:00-11:20	Scale-invariant Online Voice Activity Detection under Various Environments	Takeda, Ryu*; Komatani, Kazunori
	11:20-11:40	Sound Quality Improvement in Visual Microphone by Emphasizing Focused Area Based on Focal Rate	Nakano, Hayata*; Geng, Yuting; Iwai, Kenta; Nishiura, Takanobu
	11:40-12:00	Deep-Learning-Based Speech Enhancement with Rough-Focused Optical Laser Microphone by Reconstructing Complex Spectrum	Nakano, Yuki*; Geng, Yuting; Iwai, Kenta; Nishiura, Takanobu
	12:00-12:20	A Study on Multimodal Fusion and Layer Adapter in Emotion Recognition	Shi, Xiaohan*; Gao, Yuan; He, Jiajun; Mi, Jinyi; LI, Xingfeng; Toda, Tomoki

Session		Room	Chair
Biomedical Signal Processing and Systems		Meeting Room 8	-
Date	Time	Title	Authors
04-12-2024	11:00-11:20	Bluemarble: Bridging Latent Uncertainty in Articulatory-to-Speech Synthesis with a Learned Codebook	um, seyun*; Kim, Miseul; Kim, Doyeon; Kang, Hong-Goo
	11:20-11:40	Iterative Demographic Attentional Feature Fusion-based CNN and Transformer Network for Accurate Cuffless Blood Pressure Estimation	Tang, Liwen; Zheng, Dingchang; Chen, Fei*

	11:40-12:00	Sampling Pattern Augmentation to Enhance Deep Learning-based Image Reconstruction of MRI	Yamato, Kazuki*; Ito, Satoshi
	12:00-12:20	Data Augmentation and Assessment for Enhanced Ovarian Tumor Classification	Pham, Loan Thi*; Pham, Gia-Minh; Nguyen, Tien-Dat; Le, Hung Van; Pham, Chi-Mai; Le, Thi Lan; Vu, Duy-Hai; Vu, Hai; Tran, Thanh-Hai

Session		Room	Chair
Machine Learning and Data Analytics		Meeting Room 9	-
Date	Time	Title	Authors
04-12-2024	11:00-11:20	GMA: Green Multi-Modal Alignment for Image-Text Retrieval	Yang, Tsung-Shan*; Wang, Yun-Cheng; Wei, Chengwei; You, Suyu; Kuo, C.-C. Jay
	11:20-11:40	Improving Semi-Supervised Object Detection by ROI-Enhanced Contrastive Learning	Huang, Teng-Kuan Huang; Yeh, Mei-Chen*
	11:40-12:00	Real-time Segmentation of Coronary Artery Calcification Using Spatial Attention and Parallel Convolution	Asakawa, Tetsuya*; Hashimoto, Masashi; Miyaji, Takeshi; Shimizu, Kazuki; Nomura, Kei; Aono, Masaki

Session		Room	Chair
Speech and Language Processing		Meeting Room 10	-
Date	Time	Title	Authors

04-12-2024	11:00-11:20	LDMSE: Low Computational Cost Generative Diffusion Model for Speech Enhancement	Nishi, Yuki*; Iwano, Koji; SHINODA, Koichi
	11:20-11:40	MTFNet: Multi-Scale Transformer Framework for Robust Emotion Monitoring in Group Learning Settings	Zhang, Yi*
	11:40-12:00	Target Speaker Extraction Method by Emphasizing the Active Speech with an Additional Enhancer	Yang, Xue; Bao, Changchun*; Zhang, Xu; Chen, Xianhong

Session		Room	Chair
Multimedia Security and Forensics		Meeting Room 2	-
Date	Time	Title	Authors
04-12-2024	14:00-14:20	Compressed Deepfake Video Detection Based on 3D Spatiotemporal Trajectories	Chen, Zongmei; Liao, Xin*; Wu, Xiaoshuai; Chen, Yanxiang
	14:20-14:40	A Document Presentation Attack Detection Scheme with Optical Flow under a Flashlight	Chen, Changsheng*; Chen, Wenyu; Chen, Ximin; Li, Haodong
	14:40-15:00	Robust Image Watermarking Scheme under Halftone Distortion with Surrogate Model	Chen, Changsheng*; Li, Xijin
	15:00-15:20	A Diffusion-Based Approach for Restoring Face-swapped Images	Niu, Yuanchen; Li, Yuanman*; Zhang, Guijia; Li, Xia
	15:20-15:40	AI-generated image detectors are surprisingly easy to mislead... for now	Lyu, Zihang*; Xiao, Jun; Zhang, Cong; Lam, Kin-Man

Session	Room	Chair
---------	------	-------

Image, Video, and Multimedia		Meeting Room 3	-
Date	Time	Title	Authors
04-12-2024	14:00-14:20	Green Video Camouflaged Object Detection	Wang, Xinyu*; Chen, Hong-Shuo; Zhou, Zhiruo; You, Suyu; Madni, Azad; Kuo, C.-C. Jay
	14:20-14:40	A Survey on Objective Quality Assessment of Omnidirectional Images	Sui, Xiangjie*; Wang, Shiqi; Fang, Yuming
	14:40-15:00	Enhancing YOLOv7 with GLF-Trans for Precision in Small Object Detection	Yoshikawa, Naohito*; Ikehara, Masaaki
	15:00-15:20	Ablation Study to Derive a Computationally Efficient Deep Learning-Based Super-Resolution Approach	Jamil, Asfa*; Artusi, Alessandro
	15:20-15:40	Adaptive Spatial Re-sampling Method for Video Coding for Machines	An, Eunbin; Kim, Ayoung; Jung, Soon Heung; Choo, Hyon-Gon; Seo, Kwang-Deok*

Session		Room	Chair
Signal and Information Processing & Systems		Meeting Room 4	-
Date	Time	Title	Authors
04-12-2024	14:00-14:20	Multi-Channel Fusion Human Activity Recognition Algorithm Based on Millimeter-Wave Radar	Zhu, Junda*; Guo, Shisheng; Tang, Longzhen; Guolong, Cui

	14:20-14:40	Optimizing Computational Efficiency: In-Memory Computing with Dynamic Switching	Huang, Chao-Ting*; Tsai, Kun-Lin
	14:40-15:00	Modeling and Analysis of the Interaction between Opinions and Actions among Heterogeneous Agents	Zhang, Hangjing; Zhao, H. Vicky*
	15:00-15:20	Adaptive Subspace Clustering for Matrix Completion	Wada, Takuto*; Sasaki, Ryohei; Konishi, Katsumi
	15:20-15:40	A High-Isolation Sub-6 GHz In-Band Full-Duplex Communication System	shi, chengzhe*; Pan, Wensheng; Ma, Wanzhi; Liu, Ying; Xu, Qiang; Zhang, Zhiya; Shao, Shihai

Session		Room	Chair
Speech and Language Processing		Meeting Room 5	-
Date	Time	Title	Authors
04-12-2024	14:00-14:20	GE2E-AC: Generalized End-to-End Loss Training for Accent Classification	Watanabe, Chihiro*; Kameoka, Hirokazu
	14:20-14:40	Efficient Feature Selection for Word Embedding Dimension Reduction	Xue, Jintang*; Wang, Yun-Cheng; Wei, Chengwei; Kuo, C.-C. Jay
	14:40-15:00	Improving Speaker Consistency in Speech-to-Speech Translation Using Speaker Retention Unit-to-Mel Techniques	Zhou, Rui*
	15:00-15:20	Speech Separation using Neural Audio Codecs with Embedding Loss	Yip, Jia Qi*; Kwok, Chin Yuen; Ma, Bin; Chng, Eng Siong

	15:20-15:40	Speech Synthesis from IPA Sequences through EMA Data	Maruyama, Koki*; Sawada, Shun; Ohmura, Hidefumi; Katsurada, Kouichi
--	-------------	--	---

Session		Room	Chair
Audio Processing		Meeting Room 6	-
Date	Time	Title	Authors
04-12-2024	14:00-14:20	Learnable Cross-Correlation based Filter-and-Sum Networks for Multi-channel Speech Separation	Wang, Xianrui*; Zhang, Shiqi; He, Bo; Makino, Shoji; Chen, Jingdong
	14:20-14:40	Enhancing Neural Speech Embeddings for Generative Speech Models	Kim, Doyeon*; Song, Yanjue; Madhu, Niles; Kang, Hong-Goo
	14:40-15:00	Design of Spectrogram-Consistency Regularization Term Dependent on Observation in Independent Low-Rank Matrix Analysis for Blind Source Separation	Kojima, Takaaki*; Takamune, Norihiro; Kitamura, Daichi; Saruwatari, Hiroshi
	15:00-15:20	On Joint Dereverberation and Single Moving Source Separation with Online Source Steering	Zhang, Yiting*; Mo, Kaien; Ueda, Tetsuya; Yang, Yichen; Makino, Shoji
	15:20-15:40	New Perspectives and Insights on Distortionless Microphone Array Beamforming	Zhang, Fan*; Benesty, Jacob; Pan, Chao; Chen, Jingdong
	15:40-16:00	Block Refinement Learning for Improving Early Exit in Autoregressive ASR	Kawata, Naotaka*; Orihashi, Shota; Suzuki, Satoshi; Tanaka, Tomohiro; Ihori, Mana; Makishima, Naoki; Yamane, Taiga; Masumura, Ryo

Session		Room	Chair
Biomedical Signal Processing and Systems		Meeting Room 8	-
Date	Time	Title	Authors
04-12-2024	14:00-14:20	Postoperative Delirium Prediction Based on Preoperative Electrocardiogram and Electroencephalogram	Mito, Shogo; Miyajima, Miho; Tomioka, Hirofumi; Sato, Hitomi; Takeuchi, Takashi; Muto, Hitoshi; Kabasawa, Yuji; Harada, Hiroyuki; Eguchi, Kana; Kato, Shota; Kano, Manabu*
	14:20-14:40	A method for classification NEO-FFI answers fabricated and advantageous due to psychological bias using brainwave specific brain activity networks	ASHIKAWA, YUTO*; Ito, Takashi; Ishizu, Syohei; Kurihara, Yosuke
	14:40-15:00	Effect of White Noise on Working Memory Using Event-Related Potentials	Lee, Seung-won; LEE, Jun-Seok; Hwang, Han-Jeong*
	15:00-15:20	Automated prediction of loudness growth curve using EEG signals	Tiwari, Nitya*
	15:20-15:40	Separation of Cardiopulmonary Sound Signals for Classification of Respiratory Diseases	Zheng, Ruxin*
	15:40-16:00	Performance Improvement of Single Plane-Wave Imaging Using U-Net and Discrete Wavelet Transform	Shidara, Hiromi*; Miura, Kanta; Ishii, Takuro; Ito, Koichi; Aoki, Takafumi; Saijo, Yoshifumi ; Ohmiya, Jun

Session	Room	Chair
---------	------	-------

Best Student Paper Competition		Meeting Room 9	-
Date	Time	Title	Authors
04-12-2024	14:00-14:20	Rotation Invariant Spatio-Spectral Total Variation for Hyperspectral Image Denoising	Takemoto, Shingo*; Ono, Shunsuke
	14:20-14:40	Peer Learning via Shared Speech Representation Prediction for Target Speech Separation	Yang, Xusheng*; Zhao, Zifeng; Zou, Yuexian
	14:40-15:00	NecoBERT: Self-Supervised Learning Model Trained by Masked Language Modeling on Rich Acoustic Features Derived from Neural Audio Codec	Nakata, Wataru*; Saeki, Takaaki; Saito, Yuki; Takamichi, Shinnosuke; Saruwatari, Hiroshi
	15:00-15:20	Multi-Task Learning Approaches for Music Similarity Representation Learning Based on Individual Instrument Sounds	Imamura, Takehiro*; Hashizume, Yuka; Toda, Tomoki
	15:20-15:40	ViP-CBM: Reducing Parameters in Concept Bottleneck Models by Visual-Projected Embeddings	Qi, Ji; Wang, Huisheng; Zhao, H. Vicky*

Session		Room	Chair
Speech and Language Processing		Meeting Room 10	-
Date	Time	Title	Authors
04-12-2024	14:00-14:20	BEES: A New Acoustic Task for Blended Emotion Estimation in Speech	LI, Xingfeng*; Shi, Xiaohan; Si, Yuke; Zhang, Zilong; Cui, Feifei; Li, Yongwei; Liu, Yang; Unoki, Masashi; Akagi, Masato

	14:20-14:40	Is Corpus Truth for Human Perception?: Quality Assessment of Voice Response Timing in Conversational Corpus through Timing Replacement	Yoshikawa, Sadahiro*; Ishii, Ryo; Okada, Shogo
	14:40-15:00	Enhancing Branchformer with Dynamic Branch Merging Module for Code-Switching Speech Recognition	Hu, Hong-Jie*; Chen, Chia-Ping
	15:00-15:20	Optimizing Multi-Speaker Speech Recognition with Online Decoding and Data Augmentation Strategies	Peng, Yizhou*; Chng, Eng Siong
	15:20-15:40	Adapting OpenAI's Whisper for Speech Recognition on Code-Switch Mandarin-English SEAME and ASRU2019 Datasets	Yang, Yuhang; Peng, Yizhou*; Huang, Hao; Chng, Eng Siong; Zhong, Xionghu

Session		Room	Chair
Multimedia Security and Forensics		Meeting Room 2	-
Date	Time	Title	Authors
04-12-2024	16:20-16:40	A Study on Variable Embedding Locations of Reversible Spectral Speech Watermarking	HUANG, Xuping*; Ito, Akinori
	16:40-17:00	Normalizing Flows-Based Latent Variable Rearrangement for Generative Image Steganography	Wu, Sifan*; Dong, Li
	17:00-17:20	Detecting Spoof Voices in Asian Non-Native Speech: An Indonesian and Thai Case Study	Adila, Aulia*; Mawalim, Candy Olivia; Unoki, Masashi

	17:20-17:40	Privacy-Preserving Anomaly Detection in Bitstream Video based on Gaussian Mixture Model	Chen, Yike; Song, Yuru; Zheng, Peijia *; Du, Yusong; Luo, Weiqi
	17:40-18:00	Source Attribution for Images Generated by Diffusion-Based Text-to-Image Models: Exploring the Forensics Approach	Jiang, Xinqi; Tian, Jinyu*

Session		Room	Chair
Image, Video, and Multimedia		Meeting Room 3	-
Date	Time	Title	Authors
04-12-2024	16:20-16:40	Hyperspectral Unmixing With Row-Sparsity Enhancement: A Difference-of-Convex Approach	Naganuma, Kazuki*; Ono, Shunsuke
	16:40-17:00	How Accurate Can Large Vision Language Model Perform for Images with Compression Degradation?	Fang, Xiaohan*; CHEN, PEILIN; Wang, Meng; Wang, Shiqi
	17:00-17:20	Enhanced RefinedNet for Single Image Dehazing	Ren, Jingyu*
	17:20-17:40	Tsnake: A Time-Embedded Recurrent Contour-Based Instance Segmentation Model	Hsu, Chen-Jui; Ding, Jian-Jiun*; Shih, Chun-Jen
	17:40-18:00	A Multi-Perceptual Learning Network for Retina OCT Image Denoising and Classification	Lam, Kin-Man*

Session		Room	Chair
Signal and Information Processing & Systems		Meeting Room 4	-

Date	Time	Title	Authors
04-12-2024	16:20-16:40	Affine Combination of General Adaptive Filters	Jin, Danqi*; Chen, Yitong; Chen, Jie; Huang, Gongping
	16:40-17:00	An Annealing-Inspired Gradient-Descent Based Suboptimal Solver for Combinatorial Problems	Shu Ping, Chang; Lee, Cheng-Che; Lee, Hsin-Jung; Kuan, Chieh-Hsiung; Young, Jason Gamsun; Yao, Chia-Yu; Ding, Jian-Jiun*
	17:00-17:20	A Solution For Anomaly Detection of Red Beans In A Product Processing Line	Nguyen, Duc Hai; Do, Hiep Trong; Nguyen, Hoang-Linh-Phuong; Nguyen, Quoc-Khanh; Tran, Duc-Tan; Bui, Tien Son Tien; Nguyen, VanToi*
	17:20-17:40	A Novel kind of WVD Associated with the Linear Canonical Transform	Peng, Jia-Yin; Chen, Jian-Yi; Li, Bing-Zhao*
	17:40-18:00	A Discrete-Valued Signal Estimation by Nonconvex Enhancement of SOAV with cLiGME Model	Shoji, Satoshi*; Yata, Wataru; Kume, Keita; Yamada, Isao

Session		Room	Chair
Speech and Language Processing		Meeting Room 5	-
Date	Time	Title	Authors
04-12-2024	16:20-16:40	Frequency & Channel Attention Network for Small Footprint Noisy Spoken Keyword Spotting	Lin, Yuanxi*; Gapanyuk, Yuriy E
	16:40-17:00	Long Audio File Speaker Diarization with Feasible End-to-End Models	Huang, Kai-Wei*; Chen, Chia-Ping

	17:00-17:20	Analysis of Various Self-Supervised Learning Models for Automatic Pronunciation Assessment	Lee, Haeyoung*; Kim, Sunhee; Chung, Minhwa
	17:20-17:40	Band-Split Inter-SubNet: Band-Split with Subband Interaction for Monaural Speech Enhancement	Pan, Yen-Chou; Shen, Yih-Liang*; Liao, Yuan-Fu; Chi, Tai-Shih
	17:40-18:00	Speech Dereverberation with Deconvolution Regularized by Denoising	Hu, Haonan; Yang, Ziyue; Chen, Jie*; Zhang, Lijun

Session		Room	Chair
Audio Processing		Meeting Room 6	-
Date	Time	Title	Authors
04-12-2024	16:20-16:40	A Low-Complexity Adaptive Beamformer for Joint Reverberation and Noise Suppression	Zhang, Fan*; Pan, Chao; Chen, Jingdong; Benesty, Jacob
	16:40-17:00	Multichannel Speech Enhancement Using Complex-Valued Graph Convolutional Networks and Triple-Path Attentive Recurrent Networks	Shen, Xingyu; Zhu, Prof. Wei-Ping*
	17:00-17:20	Anomalous Machine Sound Detection Based on Time Domain Gammatone Spectrogram Feature and IDNN Model	Hafiz, Primanda Adyatma*; Mawalim, Candy Olivia; Puji Lestari, Dessi; Sakti, Sakriani; Unoki, Masashi
	17:20-17:40	Unsupervised Anomalous Sound Detection Using Timbral and Human Voice Disorder-Related Acoustic Features	Akbar Hashemi Rafsanjani, Malik*; Mawalim, Candy Olivia; Lestari, Dessi Puji; Sakti, Sakriani; Unoki, Masashi
	17:40-18:00	Real-Time Monophonic Dual-Pitch Extraction Model	Tran, Ngoc-Son; Hsieh, Pei-Chin; Shen, Yih-Liang*; Chu, Yen-Hsun; Chi, Tai-Shih

Session		Room	Chair
Biomedical Signal Processing and Systems		Meeting Room 8	-
Date	Time	Title	Authors
04-12-2024	16:20-16:40	Predictive Analysis of Driver Drowsiness Progression: Multi-Level Drowsiness Classification Using Physiological Signals	Dachoponchai, Natchira; Wongsawat, Yodchanan; Arnin, Jetsada*
	16:40-17:00	Feature Extraction for Machine Learning-based Sleep Stage Classification Using PPG-Derived Parameters and Skin Temperature	Buaruk, Suphachok; Thanaviratananich, Sikawat; Treesuthacheep, Peerasit; Deepaisarn, Somrudee*
	17:00-17:20	Parameterizing Hierarchical Particle Filters with Concept Drift for Time-varying Parameter Estimation	Murphy, Joshua*; Rosato, Conor; Millard, Andrew; Maskell, Simon
	17:20-17:40	Pop Noise Detection Using Group Delay Cepstral Coefficients	Shah, Arth Juhul*; Patil, Hemant
	17:40-18:00	Novel Estimators for the Number of Susceptible Individuals in SIR Models of Infectious Epidemics	van Wyk, Anton; McDonald, Andre M*; Rubin, David; Zhang, FangFang

Session		Room	Chair
Best Paper Competition		Meeting Room 9	-
Date	Time	Title	Authors

04-12-2024	16:20-16:40	Generalized Graph Signal Sampling under Subspace Priors by Difference-of-Convex Minimization	Yamashita, Keitaro*; Naganuma, Kazuki; Ono, Shunsuke
	16:40-17:00	Robust Adaptive Filtering Based on Adaptive Projected Subgradient Method: Moreau Enhancement of Distance Function	Sawada, Daiki; Yukawa, Masahiro*
	17:00-17:20	Fine-Grained Quantitative Emotion Editing for Speech Generation	Inoue, Sho*; Zhou, Kun; Wang, Shuai; Li, Haizhou
	17:20-17:40	Physical Domain Adversarial Attacks Against Source Printer Image Attribution	Purnekar, Nischay*; Tondi, Benedetta; Barni, Mauro
	17:40-18:00	SRC-gAudio: Sampling-Rate-Controlled Audio Generation	Li, Chenxing*; Xu, Manjie; Yu, Dong

Session		Room	Chair
Speech and Language Processing		Meeting Room 10	-
Date	Time	Title	Authors
04-12-2024	16:20-16:40	Domain Adaptation by Alternating Learning of Acoustic and Linguistic Information for Japanese Deaf and Hard-of-Hearing People	Takahashi, Kaito*; Wakabayashi, Yukoh; Ohta, Kengo; Kobayashi, Akio; Kitaoka, Norihide
	16:40-17:00	Speech emotion recognition based on crossmodal transformer and attention weight correction	Terui, Ryusei*; Yamada, Takeshi
	17:00-17:20	Unsupervised Discovery of Non-Categorical L2 Error Patterns Using Wav2Vec2.0 Code Vectors	Hong, Eunsoo*; Kim, Sunhee; Chung, Minhwa

	17:20-17:40	An Effective Contextualized Automatic Speech Recognition Approach Leveraging Self-Supervised Phoneme Features	Pai, Li-Ting*; Wang, Yi-Cheng; Yan, Bi-Cheng; Wang, Hsin-Wei; Lu, Jia-Liang; Lin, Chi-Han; Xu, Juan-Wei ; Chen, Berlin
	17:40-18:00	COIN-AT-PVAD: A Conditional Intermediate Attention PVAD	Yu, En-Lun*; Ruei-Xian, Chang; Hung, Jieh-weih; Huang, Shih-Chieh; Chen, Berlin

Session		Room	Chair
Advanced Topics on Sound Event and Scene Analysis		Meeting Room 1	-
Date	Time	Title	Authors
05-12-2024	10:20-10:40	Multi-Modal Video Summarization Based on Two-Stage Fusion of Audio, Visual, and Recognized Text Information	Yang, Zekun*; He, Jiajun; Toda, Tomoki
	10:40-11:00	Prediction-error-based Adaptive SpecAugment for Fine-tuning the Masked Model on Audio Classification Tasks	Zhang, Xiao*; XING, HAORAN; Song, Mingxue; Takeuchi, Daiki; Harada, Noboru; Makino, Shoji
	11:00-11:20	Synchronization of Signals with Sampling Rate Offset and Missing Data Using Dynamic Programming Matching	Takeuchi, Hayato*; Ono, Nobutaka
	11:20-11:40	LEAD Dataset: How Can Labels for Sound Event Detection Vary Depending on Annotators?	Koga, Naoki; Bando, Yoshiaki; Imoto, Keisuke*
	11:40-12:00	SSL-based Chewing and Swallowing Detection Using	Tsukagoshi, Toshihiro*; Koiwai, Kazuhiro; Nishida,

		Multiple Skin-contact Microphones	Masafumi; Nishimura, Masafumi
--	--	---	-------------------------------

Session		Room	Chair
Recent Advances in Multimedia Enrichment and Security		Meeting Room 2	-
Date	Time	Title	Authors
05-12-2024	10:20-10:40	Enhancing Security Using Random Binary Weights in Privacy-Preserving Federated Learning	Sawada, Hiroto*; Imaizumi, Shoko ; Kiya, Hitoshi
	10:40-11:00	Estimation of rotation angle and anisotropic scaling rate using pilot signals for watermarking	Kawano, Rinka*; Kawamura, Masaki
	11:00-11:20	On the Security of Bitstream-level JPEG Encryption with Restart Markers	Hirose, Mare*; Imaizumi, Shoko ; Kiya, Hitoshi
	11:20-11:40	Improved Ultimate Link without Markers for Projective Transformation	Yamadera, Keiji; Niimi, Michiharu*
	11:40-12:00	Detection of Diffusion-Generated Images Using Sparse Coding	Tanaka, Daishi; Niimi, Michiharu*

Session		Room	Chair
Image, Video, and Multimedia		Meeting Room 3	-
Date	Time	Title	Authors
05-12-2024	10:20-10:40	Improved Architecture for High-resolution Piano Transcription	Mi, Jinyi*; Kim, Sehun; Toda, Tomoki

		to Efficiently Capture Acoustic Characteristics of Music Signals	
	10:40-11:00	Ev3DGS:Event Enhanced 3D Gaussian Splatting from Blurry Images	Huang, Junwu; Wan, Zhexiong; Lu, Zhicheng; Zhu, Juanjuan; He, Mingyi; Dai, Yuchao*
	11:00-11:20	New Abnormal Behavior Detection for Patient Surveillance System	Han, Yujin; kim, taewan*
	11:20-11:40	Utilizing Cross Layer Attentions for Semantic Segmentation of Small Objects	Lu, Chi-Hsuan; Chung, Yu-Hsien; Cho, Jung-Hui; Yu, Chih-Chang*
	11:40-12:00	Music2Fail: Transfer Music to Failed Recorder Style	Leong, Chon In*; Chung, I-Ling; Chao, Kin Fong; Wang, Jun-You; Yang, Yi-Hsuan; Jang, Roger

Session		Room	Chair
Signal and Information Processing & Systems		Meeting Room 4	-
Date	Time	Title	Authors
05-12-2024	10:20-10:40	U-Mamba-Net: A highly efficient Mamba-based U-net style network for noisy and reverberant speech separation	Dang, Shaoxiang*; Matsumoto, Tetsuya; Takeuchi, Yoshinori; Kudo, Hiroaki
	10:40-11:00	Graph Filter Transfer for Time-Varying Signal Estimation Between Two Networks	Fukuhara, Tsutahiro*; Hara, Junya; Higashi, Hiroshi; Tanaka, Yuichi
	11:00-11:20	Few-Shot Audio Classification Model for Detecting Classroom Interactions Using LaSO Features in Prototypical Networks	Iqbal, Md Rashed*; Ritz, Christian; Yang, Jie

	11:20-11:40	Subset Random Sampling of Finite Time-vertex Graph Signals	Sheng, Hang; Shu, Qinji; FENG, HUI*; Hu, bo
	11:40-12:00	Dynamic Sensor Placement on Graphs Based on Graph Signal Sampling Theory	Nomura, Saki*; Hara, Junya; Higashi, Hiroshi; Tanaka, Yuichi

Session		Room	Chair
Speech and Language Processing		Meeting Room 5	-
Date	Time	Title	Authors
05-12-2024	10:20-10:40	Can We Estimate Purchase Intention Based on Zero-shot Speech Emotion Recognition?	Nagase, Ryotaro; Sumiyoshi, Takashi; Yamashita, Natsuo; Dohi, Kota; Kawaguchi, Yohei*
	10:40-11:00	Assessment and Improvement of Customer Service Speech with Multiple Large Language Models	Watanabe, So; Leow, Chee Siang*; Hoshino, Junichi; Utsuro, Takehito; Nishizaki, Hiromitsu
	11:00-11:20	JAM: A Unified Neural Architecture for Joint Multi-granularity Pronunciation Assessment and Phone-level Mispronunciation Detection and Diagnosis Towards a Comprehensive CAPT System	He, Yue-Yang*; Yan, Bi-Cheng; Lo, Tien-Hong; Lin, Meng-Shin; Hsu, Yung-Chang; Chen, Berlin
	11:20-11:40	Data Augmentation Methods and Influence of Speech Recognition Performance for TED Talk's English to Japanese Speech Translation	Masuda, Kento*; Yamamoto, Kazumasa; nakagawa, seiichi
	11:40-12:00	Empower Typed Descriptions by Large Language Models for Speech Emotion Recognition	Wu, Haibin; Chou, Huang-Cheng*; Chang, Kai-Wei; Goncalves, Lucas; Du, Jiawei; Jang,

			Jyh-Shing Roger; Lee, Chi-Chun; Lee, Hung-yi
--	--	--	--

Session		Room	Chair
Audio Processing		Meeting Room 6	-
Date	Time	Title	Authors
05-12-2024	10:20-10:40	Wind Noise Reduction with Orthogonal Polynomial Expansion	Du, Li*; Zhang, Lijun
	10:40-11:00	Few-Shot Open-Set Keyword Spotting with Multi-Stage Training	Li, LoYa*; Lo, Tien-Hong; Hung, Jieih-weih; Huang, Shih-Chieh; Chen, Berlin
	11:00-11:20	Self-Supervised Augmented Diffusion Model for Anomalous Sound Detection	Yin, Jiawei; gao, yu*; Zhang, Wenbin; Zhang, Mingjun
	11:20-11:40	Murmur Separation and Classification from Heart Sound Using Constrained Singular Spectrum Analysis and Wavelet Transform	Qi, Yuanyang*; Sanei, Saeid
	11:40-12:00	A Non-Intrusive Speech Quality Assessment Model using Whisper and Multi-Head Attention	Lin, Guojian; Tsao, Yu; Chen, Fei*

Session		Room	Chair
Emerging Technologies and Applications Of Image Processing And Computer Vision		Meeting Room 9	-
Date	Time	Title	Authors

05-12-2024	10:20-10:40	Confidence-Aware Learning for Person Re-identification with Noisy Labels	Kim, Duhyun*; Sim, Jae-Young
	10:40-11:00	Test-Time Optimization for Post-Processing of Compressed Videos	Kim, Hongil; Han, Changwoo; Kim, Donghyun; Lim, Sung-Chang; Jung, Seung-Won*
	11:00-11:20	Lifelong Person Re-Identification with Backward-Compatibility	Oh, Minyoung; Sim, Jae-Young*
	11:20-11:40	Enhancing Semiconductor X-RAY Images: A Framework Combining Denoising and Super-Resolution Modules With a Novel Dataset	Shim, Jae Hoon*; Kim, Min Woo; Lee, Sang Hwa; Cho, Nam Ik
	11:40-12:00	Monocular Depth Estimation for Autonomous Driving Based on Instance Clustering Guidance	Kim, Dahyun*; Jin, Dongkwon; Kim, Chang-Su

Session		Room	Chair
Advanced Signal Processing for Information Collection and Data Analysis in Wireless Environmental Sensing		Meeting Room 10	-
Date	Time	Title	Authors
05-12-2024	10:20-10:40	Data-Driven Tuning for Weighted Least Square of BLE-AoA-based Indoor Localization	Ohashi, Ginji; Ibi, Shinsuke*; Takahashi, Takumi; Iwai, Hisato
	10:40-11:00	Observation of the terrestrial radio environment using the	Obata, Takatoshi*; Takyu, Osamu; Inage, Kei; Fujii, Takeo;

		low earth orbit satellite constellation	Yoshida, Kohei; Ariyoshi, Masayuki
	11:00-11:20	Deep Unfolding Aided Parameter Optimization for Multi-task Diffusion LMS Algorithm	Tong, Xiaoqing*; Hayashi, Kazunori
	11:20-11:40	Reduced-dimensional MUSIC Algorithm for Frequency Diverse Array in MIMO Radar System	Zhu, Beizuo*; Hayashi, Kazunori; Mori, Hiroki
	11:40-12:00	Collection of Correlated Information from Superimposed Multiple Chirp Signals	Aoyama, Koki*; Adachi, Koichi

Session		Room	Chair
New Frontiers in Biometric Authentication		Meeting Room 1	-
Date	Time	Title	Authors
05-12-2024	14:00-14:20	A Quasilinear-Time CVP Algorithm for Triangular Lattice Based Fuzzy Extractors and Fuzzy Signatures	Takahashi, Kenta*; Nakamura, Wataru
	14:20-14:40	Enhancing Remote Adversarial Patch Attacks on Face Detectors with Tiling and Scaling	Okano, Masora*; Ito, Koichi; Nishigaki, Masakatsu; Ohki, Tetsushi
	14:40-15:00	Multibiometrics Using a Single Face Image	Ito, Koichi*; Tonosaki, Taito; Aoki, Takafumi; Ohki, Tetsushi; Nishigaki, Masakatsu

	15:00-15:20	Multi-Observed Authentication: A secure and usable authentication based on multi-point observation of a single physical credential	Hatakeyama, Wataru*; Nozaki, Shinnosuke; Serizawa, Ayumi; Yoshirira, Mizuho; Fujita, Masahiro; Yoshimura, Ayako; Ohki, Tetsushi; Nishigaki, Masakatsu
--	-------------	--	---

Session		Room	Chair
Recent Advances in Multimedia Enrichment and Security		Meeting Room 2	-
Date	Time	Title	Authors
05-12-2024	14:00-14:20	Generation of Target Speech with Speaker Individuality Based on Accent Conversion for English Pronunciation Learning	Hamakawa, Rei; Niimi, Michiharu*
	14:20-14:40	Proposal of Blind Extractable Additive Video Watermarking Method	Harada, Nao*; Kawano, Rinka; Kawamura, Masaki
	14:40-15:00	Transfer-Based Adversarial Attack Against Multimodal Models by Exploiting Perturbed Attention Region	Disabato, Raffaele*; Maung Maung, April Pyone; Nguyen, Huy Hong; Echizen, Isao
	15:00-15:20	A Permutation-based Reversible Data Hiding Method with Zero Visual Distortion	Zhu, Wendi*; Wong, KokSheik; Kuribayashi, Minoru

Session		Room	Chair
Image, Video, and Multimedia		Meeting Room 3	-
Date	Time	Title	Authors

05-12-2024	14:00-14:20	VietSing: A High-quality Vietnamese Singing Voice Corpus	Vu, Minh Duc*; Wei, Zhou; Bhattarai, Binit; Teh, Kah Kuan; Dat, Tran Huy
	14:20-14:40	Inertial Strengthened CLIP model for Zero-shot Multimodal Egocentric Activity Recognition	He, Mingzhou; Wang, Haojie; Zhou, Shuchang; Wu, Qingbo*; Ngan, King Ngi; Meng, Fanman; Li, Hongliang
	14:40-15:00	Optimization of the Intensity Aware Loss for Dynamic Facial Expression Recognition	Lau, Davy Tec-Hinh; Ding, Jian-Jiun*; Muller, Guillaume
	15:00-15:20	Dictionary Learning Based Two-stage Near-lossless Video Compression	Zhang, Zuhai; Jia, Luheng*; Song, Li; Zhu, Shuyuan; Guo, Yuanfang; Jia, Kebin

Session		Room	Chair
Signal and Information Processing & Systems		Meeting Room 4	-
Date	Time	Title	Authors
05-12-2024	14:00-14:20	Dictionary Learning for Directed Graph Signals via Augmented GFT	Naito, Tsubasa*; Ito, Ryuto; Tanaka, Yuichi; Muramatsu, Shogo
	14:20-14:40	Robust Quantile Regression Under Unreliable Data	Shoji, Yoshifumi*; Yukawa, Masahiro
	14:40-15:00	Ensemble learning based head-related transfer function personalization using anthropometric features	Shen, Yih-Liang*; Chi, Tai-Shih
	15:00-15:20	Blind Estimation of Room Volume from Reverberant Speech Based on the Modulation Transfer Function	Siripool, Nutchanon*; kongprawechnon, Waree; Unoki, Masashi

Session		Room	Chair
Speech and Language Processing		Meeting Room 5	-
Date	Time	Title	Authors
05-12-2024	14:00-14:20	Disentangling Speaker Representations from Intuitive Prosodic Features for Speaker-Adaptative and Prosody-Controllable Speech Synthesis	Pengyu, Cheng*
	14:20-14:40	A Pilot Study of Applying Sequence-to-Sequence Voice Conversion to Evaluate the Intelligibility of L2 Speech Using a Native Speaker's Shadowings	Geng, Haopeng *; Saito, Daisuke; Minematsu, Nobuaki; Geng, Haopeng
	14:40-15:00	EADSum: Element-Aware Distillation for Enhancing Low-Resource Abstractive Summarization	Lu, Jia-Liang*; Yan, Bi-Cheng; Wang, Yi-Cheng; Lo, Tien-Hong; Wang, Hsin-Wei; Pai, Li-Ting; Chen, Berlin
	15:00-15:20	A Tiny Whisper-SER: Unifying Automatic Speech Recognition and Multi-label Speech Emotion Recognition Tasks	Chou, Huang-Cheng*

Session		Room	Chair
Audio Processing		Meeting Room 6	-
Date	Time	Title	Authors
05-12-2024	14:00-14:20	EEND-EM: End-to-End Neural Speaker Diarization with EM-Network	Woo, Beom Jun*; Yoon, Ji Won; Han, Min Hyun; Moon, Chan Yeong; Kim, Nam Soo

	14:20-14:40	Personal Voice Activity Detection With Ultra-Short Reference Speech	Xu, Longting; Zhang, Mingjun; Zhang, Wenbin; Wang, Tianyi; Yin, Jiawei; gao, yu*
	14:40-15:00	An Investigation on the Speech Recovery from EEG Signals Using Transformer	Mizuno, Tomoaki*; Kishida, Takuya; Yoshimura, Natsue; Nakashika, Toru

Session		Room	Chair
Audio Processing		Meeting Room 8	-
Date	Time	Title	Authors
05-12-2024	14:00-14:20	WavLM and Omni-Scale CNNs: Enhancing Boundary Detection in Partially Spoofed Audio	Li, Menghan*; Huang, Zhihua
	14:20-14:40	Semi-Supervised Far-Field Speaker Verification with Distance Metric Domain Adaptation	Wang, Han*; He, Mingrui; Zhang, Mingjun; Xu, Longting
	14:40-15:00	Non-Target Conversion Based Speech Steganography for Secure Speech Communication System	Zhang, Mingjun; Feng, Yan; gao, yu; Xu, Longting*
	15:00-15:20	Enhancing Acoustic Scene Classification with Layer-wise Fine-Tuning on the SSAST Model	Hao, Shuting*; Saito, Daisuke; Minematsu, Nobuaki

Session		Room	Chair
High Performance Image and Video Processing and Applications		Meeting Room 9	-
Date	Time	Title	Authors

05-12-2024	14:00-14:20	Forward Prediction-Guided Cross-Partition Targeted Pruning for VVenC	Tang, Jingyuan*; Sun, Songlin
	14:20-14:40	Contrastive Learning Based Knowledge Distillation for Enhancing Defect Detection	Guo, Jing-Ming; Yuan, Lun-Da; HUANG, CIAN*; Zeng, Yi-Chong
	14:40-15:00	Screen Content Encoding Network Based on Deep Contextual Information	Gong, Tianyu*; Zhang, Tao; Zhong, Ye; Zhang, Mengmeng; Bai, Huihui
	15:00-15:20	A Coarse-to-Fine Change Detection Framework for Remote Sensing Sparse Cultivated Land	hu, yuan*; Zhang, Yifan; Ma, Mingyang; Mei, Shaohui

Session		Room	Chair
Advancements in Biosignal Decoding and Neuromodulation for Human Function Enhancement		Meeting Room 10	-
Date	Time	Title	Authors
05-12-2024	14:00-14:20	Context-FFT: A Context Feed Forward Transformer Network for EEG-based Speech Envelope Decoding	Chen, Ximin; Ding, Yuting; Yan, Nan; Chen, Changsheng; Chen, Fei*
	14:20-14:40	Effect of Dynamic Binaural Beats on Concentration Enhancement	LEE, Jun-Seok; Lee, Yun-Sung; Hwang, Han-Jeong*
	14:40-15:00	EEG-based Evaluation of Enjoyment Emotion during cognitive-motor task	Aoki, Haruna*; Zhang, Sinan; Ono, Yumie
	15:00-15:20	Exploring Brain Connectivity Patterns and Cognitive Resilience in	ks, Kapeleshh*; Wei, Chen; Domer, Prince Aldrin; Ji, Hong

		Aging: A Study with the LEMON Dataset	
	15:20-15:40	A Study on Packet-Level Index Modulation Using Frequency Offsets within a LoRaWAN Channel	ohta, mai*; Matsuura, Hiroki; Fujii, Takeo

Session		Room	Chair
Wireless Communications and Networking		Meeting Room 1	-
Date	Time	Title	Authors
05-12-2024	16:40-17:00	Blind Self-Interference Analog Canceller with Differential Delay for Backscatter Communications	Nishikawa, Koichi; Ibi, Shinsuke*; Takahashi, Takumi; Iwai, Hisato
	17:00-17:20	IoT-based Smart Attendance System using Face Recognition and Motion Detection	Saadon, Umi Syamimi*; Lim, Chern Hong

Session		Room	Chair
Recent Advances in Multimedia Enrichment and Security		Meeting Room 2	-
Date	Time	Title	Authors
05-12-2024	16:40-17:00	Generation of Photo Slideshow with Song based on Closeness between Concept of Lyrics and That of Images	Hashimoto, Mei; Niimi, Michiharu*
	17:00-17:20	Disposable-key-based image encryption for collaborative	Aso, Rei*; Shiota, Sayaka; Kiya, Hitoshi

		learning of Vision Transformer	
	17:20-17:40	Significance of Lower Frequency Regions for Audio Deepfake Detection	Shah, Arth Juhul*; Patil, Hemant
	17:40-18:00	EAViT: External Attention Vision Transformer for Audio Classification	Iqbal, Aquib; Zim, Abid Hasan; Tonmoy, Md Asaduzzaman; Zhou, Limengnan ; Malik, Asad*; Kuribayashi, Minoru

Session		Room	Chair
Image, Video, and Multimedia		Meeting Room 3	-
Date	Time	Title	Authors
05-12-2024	16:40-17:00	A Two-Stage Method for 3D Architecture Wireframe Reconstruction from Airborne LiDAR Point Cloud	Zhang, Jiahao; Liu, Qi*; Hui, Le; Dai, Yuchao
	17:00-17:20	Secure Moving Object Detection Transformer in Compressed Video with Feature Fusion	Song, Yuru; Chen, Yike; Zheng, Peijia *; Du, Yusong; Luo, Weiqi
	17:20-17:40	NeRF-FCM: Attention-based Feature Calibration Mechanisms for 3D Object Detection Using NeRF	Goshu, Hana Lebeta*; Xiao, Jun; Chan, Kin-Chung; Zhang, Cong; Gameda, Mulugeta Tegegn; Lam, Kin-Man
	17:40-18:00	High-Quality Facial Pose Generation with Latent Space Processing	Siu, Wan-Chi*; Cheng, Wing-Ho; Chan, H Anthony

Session		Room	Chair
---------	--	------	-------

Signal and Information Processing & Systems		Meeting Room 4	-
Date	Time	Title	Authors
05-12-2024	16:40-17:00	Significance of Entropy Based Features For Dysarthric Severity Level Classification	Avula, Meghana*; Pusuluri, Aditya; Patil, Hemant
	17:00-17:20	Incorporating Auditory Processing into Undergraduate Signal Processing Courses to Enhance Student Learning	Nie, Kaibao *
	17:20-17:40	A Real-Time Platform for Portable and Scalable Active Noise Mitigation for Construction Machinery	Peksi, Santi; Gan, Woon Seng*; Lai, Chung Kwan; Lee, Yen Theng; Shi, Dongyuan; Lam, Bhan

Session		Room	Chair
Speech and Language Processing		Meeting Room 5	-
Date	Time	Title	Authors
05-12-2024	16:40-17:00	A Comparative Study on the Biases of Age, Gender, Dialects, and L2 speakers of Automatic Speech Recognition for Korean Language	Na, Jonghwan; Park, Yeseul; Lee, Bowon*
	17:00-17:20	Targeted Representation with Information Disentanglement Encoding Networks in Tasks	Nagawaki, Takumi*; Ikeda, Keisuke; Tamura, Satoshi; Chike, Kohei; Nagano, Hiroyuki; Nose, Masaki
	17:20-17:40	PG-MDD: Prompt-Guided Mispronunciation Detection and Diagnosis Leveraging Articulatory Features	Lin, Meng-Shin*; Yan, Bi-Cheng; Lo, Tien-Hong; Wang, Hsin-Wei; He,

			Yue-Yang; Chao, Wei-Cheng; Chen, Berlin
--	--	--	--

Session		Room	Chair
Audio Processing		Meeting Room 6	-
Date	Time	Title	Authors
05-12 -2024	16:40- 17:00	Experimental Evaluation of Speech Enhancement for In-Car Environment Using Blind Source Separation and DNN-based Noise Suppression	Takeuchi, Yutsuki*; Nakashima, Taishi; Ono, Nobutaka; Takazawa, Takashi; Shimano, Shuhei; Tsuchiya, Yoshinori
	17:00- 17:20	Auxiliary-Function-Based Steering Vector Estimation Method for Spatially Regularized Independent Low-Rank Matrix Analysis	Hirata, Sota*; Takamune, Norihiro; Yamaoka, Kouei; Kitamura, Daichi; Saruwatari, Hiroshi; Takahashi, Yu; KONDO, Kazunobu
	17:20- 17:40	Two-stage Framework for Robust Speech Emotion Recognition Using Target Speaker Extraction in Human Speech Noise Conditions	Mi, Jinyi*; Shi, Xiaohan; Ma, Ding; He, Jiajun; Fujimura, Takuya; Toda, Tomoki
	17:40- 18:00	Data generation for speaker diarization by speaker transition information	Ichikawa, Keigo*; Ueno, Sei; Lee, Akinobu

Session		Room	Chair
Audio Processing		Meeting Room 8	-
Date	Time	Title	Authors
05-12 -2024	16:40- 17:00	Generating Room Impulse Responses Using Neural Networks Trained with Weighted Combinations of Acoustic Parameter Loss Functions	Ren, Hualin*; Ritz, Christian; Zhao, Jiahong; Zheng,

			Xiguang; Jang, Daeyoung
	17:00-17:20	Audio Similarity Detection	Malhotra, Siddharth; Mankad, Sapan H*
	17:20-17:40	Towards a B-format Ambisonic Room Impulse Response Generator Using Conditional Generative Adversarial Network	Ren, Hualin*; Ritz, Christian; Zhao, Jiahong; Zheng, Xiguang; Jang, Daeyoung
	17:40-18:00	What to Refer and How? - Exploring Handling of Auxiliary Information in Target Speaker Extraction	Hayashi, Tomohiro*; Ogino, Riku; Saijo, Kohei; Ogawa, Tetsuji

Session		Room	Chair
High Performance Image and Video Processing and Applications		Meeting Room 9	-
Date	Time	Title	Authors
05-12-2024	16:40-17:00	Efficient Adaptation for Real-World Omnidirectional Image Super-Resolution	Yang, Cuixin*; Dong, Rongkang; Lam, Kin-Man
	17:00-17:20	More Direct and stage-wise network for Face Super Resolution	Horiguchi, Yohei*
	17:20-17:40	Camera Focal Length Prediction for Neural Novel View Synthesis from Monocular Video	Chakraborty, Dipanita*; Chiracharit, Werapon; Chamnongthai, Kosin; Okada, Minoru
	17:40-18:00	Scene-Segmentation-Based Exposure Compensation for Tone Mapping of High Dynamic Range Scenes	Kinoshita, Yuma*; Kiya, Hitoshi

Session		Room	Chair
Advancements in Biosignal Decoding and Neuromodulation for Human Function Enhancement		Meeting Room 10	-
Date	Time	Title	Authors
05-12-2024	16:40-17:00	Effect of Phase-Locked Transcranial Alternating Current Stimulation on Vocal tremor	WANG, JUNTING*; Koganemaru, Satoko; Shima, Atsushi; Cao, Yedi; Hirakawa, Kana; Iwagana, Ken; Suehiro, Atsushi; Maekawa, Keiko; Mima, Tatsuya; Ono, Yumie
	17:00-17:20	Complex CNN incorporating Hilbert transform for steady-state visual evoked potential BCI	Takata, Rintaro*; Washizawa, Yoshikazu
	17:20-17:40	Electroencephalogram-Based Effective Features for Sustained Attention Assessment in Conversation	Togashi, Masaya; Chanpornpakdi, Ingon; Tanaka, Toshihisa*

Session		Room	Chair
Embedded and Real-Time Systems for AI and Signal Processing Applications		Meeting Room 1	-
Date	Time	Title	Authors
06-12-2024	09:00-09:20	Accelerated Real-Time Local Maxima Detection in Video Streams Using FPGA Technology	Nayazirly, Anindhita; Salomo, Yahwista*; Adiono, Trio; Syafalni,

			Infall; Sutisna, Nana; Mulyawan, Rahmat
	09:20-09:40	A Configurable OFDM Baseband Processor for RF-UOWC System-on-Chip	Adiono, Trio; Setiawan, Erwin*; Jonathan, Michael; Mulyawan, Rahmat; Sutisna, Nana; Syafalni, Infall; Popoola, Wasiu
	09:40-10:00	Hammering Sound Inspection System Using HPSS and Gradient Boosting with a Wall-Climbing Robot	Koyama, Nichika*
	10:00-10:20	Implementation of Real Time Oscillometric Based Algorithm for Blood Pressure Measurement in Patient Monitor	Adiono, Trio; Amadeus, Clarence*; Thomi, Teuku Rafifsyah; Sinaga, Sindy Novaria Cicilya

Session		Room	Chair
Selected Papers from APSIPA Workshop on Advanced Signal and Information Processing		Meeting Room 2	-
Date	Time	Title	Authors
06-12-2024	09:00-09:20	Automated Pseudo-Label Generation and Parallel Computing for Enhanced Few-Shot Medical Image Segmentation	Do, Ha Thanh *; Nguyen Trong, Duc; Do, Tien-Dung
	09:20-09:40	Enhanced Sparse Convolutional Detection Model for 3D Object Detection in Autonomous Vehicles Adapted to Traffic Conditions in Vietnam	Do, Ha Thanh *; Dung, Vu Hoang; Nguyen, Kien Trung

	09:40-10:00	Enhancing Cell Segmentation using Deep Learning Models by Custom Processing Techniques	Do, Ha Thanh *; Nguyen, Van De; Dang Hoang, Minh Huong; Huy, Nguyễn Quang; Dinh Manh, Cuong Initail
	10:00-10:20	Marker-Aware Ovarian Tumor Segmentation from Ultrasound Images	Bui, Hoang-Son*; Tran, Sy-Hoang; Nguyen, Thuy-Binh; Tran, Thanh-Hai; Vu, Hai; Lan, Le Thi

Session		Room	Chair
Image, Video, and Multimedia		Meeting Room 3	-
Date	Time	Title	Authors
06-12-2024	09:00-09:20	ACE-Flow: Auto Color Encoding for Enhanced Low-Light Image Restoration	Qiu, Jiachen; Zuo, Yushen; Lam, Kin-Man*
	09:20-09:40	PBJDT: Point-Based Joint Detection-and-Tracking	Lee, Zhen-Xun; Ding, Jian-Jiun*
	09:40-10:00	Capturing Dynamic Identity Features for Speaker-Adaptive Visual Speech Recognition	Kashiwagi, Sara*; Tanaka, Keitaro; Morishima, Shigeo
	10:00-10:20	A Byte-based GPT-2 Model for Bit-flip JPEG Bitstream Restoration	Qin, Hao; SUN, Haoran; Wang, Yi*

Session		Room	Chair
Acoustic Scene Analysis and Signal Enhancement Based on Advanced Signal		Meeting Room 4	-

Processing and Machine Learning			
Date	Time	Title	Authors
06-12-2024	09:00-09:20	Successive Speaker Relative Transfer Function Estimation Through Relative Transfer Matrix in Noisy Reverberant Environments	Manamperi, Wageesha*; Abhayapala, Thushara
	09:20-09:40	Heavy-tailed Distributions-Based Online Semi-blind Source Separation for Nonlinear Echo Cancellation	Zhang, Liyuan*; Wang, Xianrui; Yang, Yichen; Ueda, Tetsuya; Makino, Shoji; Chen, Jingdong
	09:40-10:00	A Single-InputBinaural-Output Perceptual Rendering Based Speech Separation Method in Noisy Environments	zheng, tianqin*; Pei, Hanchen; Pan, Ningning; Jin, Jilu; Huang, Gongping; Chen, Jingdong; Benesty, Jacob
	10:00-10:20	Real-Time Noise Estimation for Lombard-Effect Speech Synthesis in Human--Avatar Dialogue Systems	Ishikawa, Yuto*; Take, Osamu; Nakamura, Tomohiko; Takamune, Norihiro; Saito, Yuki; Takamichi, Shinnosuke; Saruwatari, Hiroshi

Session		Room	Chair
Speech and Language Processing		Meeting Room 5	-
Date	Time	Title	Authors
06-12-2024	09:00-09:20	EMO-Codec: An In-Depth Look at Emotion Preservation Capacity of Legacy and Neural Codec Models	Ren, Wenzhe*; Lin, Yi-Cheng; Chou, Huang-Cheng; Wu, Haibin; Wu, Yi-Chiao; Lee,

		With Subjective and Objective Evaluations	Hung-yi; Lee, Chi-Chun; Wang, Hsin-Min; Tsao, Yu
	09:20-09:40	Analytic Study of Text-Free Speech Synthesis for Raw Audio using a Self-Supervised Learning Model	Park, Joonyong*; Saito, Daisuke; Minematsu, Nobuaki
	09:40-10:00	Investigating the Language Independence of Voice Activity Projection Models through Standardization of Speech Segmentation Labels	Sato, Yuki*; Chiba, Yuya; Higashinaka, Ryuichiro
	10:00-10:20	A Preliminary Study on Analysing Mandarin Tone Values of Romance L2 Mandarin Learners	Li, Wu-Hao*; Liu, Te-hsin; CHIANG, Chen Yu

Session		Room	Chair
Signal Processing for Drone Audition & Recent Advances in Intelligent Signal Processing		Meeting Room 6	-
Date	Time	Title	Authors
06-12-2024	09:00-09:20	Relative Transfer Matrix for Drone Audition Applications: Source Enhancement	Manamperi, Wageesha*; Abhayapala, Thushara
	09:20-09:40	Beamforming informed independent low-rank matrix analysis for sound source enhancement in unmanned aerial vehicles	Teh, Jin Xuan*; Takamune, Norihiro; Saruwatari, Hiroshi; Yen, Benjamin; Kingan, Michael; Hioka, Yusuke
	09:40-10:00	SMoLnet-T: An Efficient Complex-spectral Mapping Speech Enhancement Approach with Frame-wise CNN and	Tan, Zhi-Wei*; Khong, Andy W H

		<u>Spectral Combination Transformer for Drone Audition</u>	
	10:00-10:20	<u>Integrating VGGSK and BEATs for Enhanced Sound Event Detection: A Semi-Supervised GRU-Based System with Weak Labels and Synthetic Soundscapes</u>	Chan, Po-Cheng*; Chen, Wei-Yu; Wang, Jia-Ching; Lu, Chung-li; Chuang, Hsiang Feng; cheng, yu-han
	10:20-10:40	<u>Drone audition: implementation of an indoor multi-drone system for sound source tracking</u>	Yen, Benjamin*; Nakadai, Kazuhiro

Session		Room	Chair
Converging AI and Computer Vision: Innovations and Potential		Meeting Room 8	-
Date	Time	Title	Authors
06-12-2024	09:00-09:20	<u>Hyperspectral Anomaly Detection Using Robust Principal Component Analysis with Autoencoding Adversarial Networks</u>	Emoto, Atsuya; Matsuoka, Ryo*
	09:20-09:40	<u>Optimising Neural Networks with Fine-Grained Forward-Forward Algorithm: A Novel Backpropagation-Free Training Algorithm</u>	Gong, James; Li, Bruce; Abdulla, Waleed*
	09:40-10:00	<u>Two-Way Malaysian Sign Language Communication System for Inclusive Education</u>	HII, Veron Zhen Liang; LO, Aaron Ken Kiat; LEE, Ida Pei Xin; ABUAN, ALEC VINCE GONZALES; Lee, Sue Han*; Then, Patrick HangHui

	10:00-10:20	PRTGaussian: Efficient Relighting Using 3D Gaussians with Precomputed Radiance Transfer	Zhang, Libo*; Han, Yuxuan; Lin, Wenbin; Ling, Jingwang; Xu, Feng
--	-------------	---	--

Session		Room	Chair
AI-Driven Innovations in Cybersecurity Advanced Applications in Signal Processing, Multimedia Security, and Privacy		Meeting Room 9	-
Date	Time	Title	Authors
06-12-2024	09:00-09:20	ET-SSM: Linear-Time Encrypted Traffic Classification Method Based On Structured State Space Model	YanJun, Li*; Zhao, Xiangyu; Zhengpeng, Zha; Ling, Zhen-Hua
	09:20-09:40	Toward Universal Detector for Synthesized Images by Estimating Generative AI Models	Seo, Ryota*; Kuribayashi, Minoru; Ura, Akinobu; Mallet, Antoine; Coganne, Rémi; Mazurczyk, Wojciech; Megías, David
	09:40-10:00	Innovative Information Hiding in H.266/VVC Using Sub-Block Transform Technique	Hau, Joan*; Tew, Yiqi; Tan, Li Peng
	10:00-10:20	GGMDDC: An Audio Deepfake Detection Multilingual Dataset	Purohit, Ravindrakumar M.*; Shah, Arth Juhul; Patil, Hemant

Session	Room	Chair
Embedded and Real-Time Systems for AI and Signal	Meeting Room 1	-

Processing Applications			
Date	Time	Title	Authors
06-12-2024	10:40-11:00	Exploration Robot Based On YOLOv8 Algorithm	Syafalni, Infall*; Winasta Sinisuka, Angelica; Kalam Amal Tauhid, Dwi; Ahmad, Farrel; Alif Putra Yasa, Muhammad; Alexander Wen, Steven; Setiawan, Erwin; Sutisna, Nana; Adiono, Trio
	11:00-11:20	Optimizing Deep Q-Network for Shortest Path Computation of Mobile Robot Agents	Sumarudin, A*; Sutisna, Nana; Syafalni, Infall; Riyanto Trilaksono, Bambang; Adiono, Trio
	11:20-11:40	Leveraging IoT and Machine Learning for Efficient Rice Stock Monitoring and Prediction	Sutisna, Nana*; Prawira Nugroho, Aditya; Jeffrey, Christopher; Ramadhana, Rizky; Mahendra, Ronggur; Jonathan, Michael; Syafalni, Infall; Adiono, Trio
	11:40-12:00	Comparative Evaluation of Fine-Tuned Hybrid Transformer and Band-Split Recurrent Neural Networks for Music Source Separation	Kalang Al Qalyubi, Ken; Ahmadi, Nur*; Puji Lestari, Dessi

Session		Room	Chair
Selected Papers from APSIPA Workshop on Advanced Signal and Information Processing		Meeting Room 2	-
Date	Time	Title	Authors

06-12-2024	10:40-11:00	Enhancing Shear Wave Propagation Analysis in Tissue with Directional Filtering of Reflected Waves	Luong, Hai Quang*; Tran, Nghia Duc; Nguyen, Hiep; Sinh Cong, Lam; Tran, Duc-Tan
	11:00-11:20	Structural Analysis of Asian and African Rice Panicles via Transfer Learning	Dinh, Tran Hiep*
	11:20-11:40	New approach for Alzheimer's disease classification using topographic maps and deep learning model	Le, Quoc Anh*; Thinh, Nguyen hong
	11:40-12:00	M-IRRA: A Multilingual Model for Text-based Person Search	Tran, Phong Ngoc Hung; Phan, Thi-Hoai; Nguyen, Thuy-Binh; Do, Ngoc-Diep; Nguyễn, Quân Hồng; Tran, Thanh-Hai ; Duong, Thanh Thi-Hien; Le, Thi Lan*

Session		Room	Chair
Image, Video, and Multimedia		Meeting Room 3	-
Date	Time	Title	Authors
06-12-2024	10:40-11:00	GMNER-LF: Generative Multi-modal Named Entity Recognition Based on LLM with Information Fusion	Hu, Huiyun*; Kong, Junda; Xiao, Bo; Wang, Fei; Ge, Yang; Sun, Hongzhi
	11:00-11:20	WildPose: HRNet-based Lightweight and Efficient Wildlife Pose Estimation	BAKANA, SIBUSISO R*; Zhang, Yongfei ; Twala, Bhekisipho
	11:20-11:40	New approach on Smiling faces with Domain Transfer in Latent Space	Siu, Wan-Chi*; DUAN, Mingfei; Hui, Chun Chuen

Session		Room	Chair
Advanced Topics for Automatic Speakers Recognition		Meeting Room 4	-
Date	Time	Title	Authors
06-12-2024	10:40-11:00	JOSEPH: PHONETIC-AWARE SPEAKER EMBEDDING FOR FAR-FIELD SPEAKER VERIFICATION	JIN, Zezhong*; TU, Youzhi; Mak, Manwai
	11:00-11:20	Vocal Tract Length Perturbation-based Pseudo-Speaker Augmentation Considering Speaker Variability for Speaker Verification	Zou, Hengyi*; Shiota, Sayaka
	11:20-11:40	Differences Between Singer and Speaker Verification: Training Singer Feature Representation Extractor Utilizing Singing Voice Characteristics	Toma, Sayaka*; Ariga, Tomoki; Higuchi, Yosuke; Hayasaka, Ichiju; Shigyo, Rie; Ogawa, Tetsuji

Session		Room	Chair
Speech and Language Processing		Meeting Room 5	-
Date	Time	Title	Authors
06-12-2024	10:40-11:00	Developing a Multilingual Spontaneous L2 Speech Corpus for Automated Proficiency Assessment	Han, Seunghee*; Kim, Sunhee; Chung, Minhwa
	11:00-11:20	Prediction of Negative User Reactions Towards System Responses During Attentive Listening	Lala, Divesh*; Inoue, Koji; Kawahara, Tatsuya

	11:20-11:40	Data Selection using Spoken Language Identification for Low-Resource and Zero-Resource Speech Recognition	Chen, Jianan*; Chu, Chenhui; Li, Sheng; Kawahara, Tatsuya
--	-------------	---	---

Session		Room	Chair
Signal Processing for Drone Audition & Recent Advances in Intelligent Signal Processing		Meeting Room 6	-
Date	Time	Title	Authors
06-12-2024	10:40-11:00	Drone audition: dataset and methods for ground surface material classification using drone noise in outdoor environment	Yano, Tsubasa*; Yen, Benjamin; Nakadai, Kazuhiro
	11:00-11:20	Seismic-ionospheric Precursor Prediction Using Deep Learning	Pham, Tung Bach*; Chang, Pao-Chi; Wang, Jia-Ching
	11:20-11:40	Swarm Active Audition System with Robots and Drones for a Search and Rescue Task	Nakadai, Kazuhiro*; Kumon, Makoto; Sasaki, Yoko; Hoshiba, Kotaro; Yen, Benjamin
	11:40-12:00	Implementation of a Robot Operation System-based network for sound source localization using multiple drones	Yamamoto, Takumi*; Hoshiba, Kotaro; Yen, Benjamin; Nakadai, Kazuhiro

Session		Room	Chair
Converging AI and Computer Vision:		Meeting Room 8	-

Innovations and Potential			
Date	Time	Title	Authors
06-12-2024	10:40-11:00	RepViT Based Lightweight Architecture for Distracted Driving Detection	Jian, Muwei*; Ling, Yukun
	11:00-11:20	HSIC as Information Compression for Training Deep Neural Network	Sofi, Roshan Birjais*; Wang, Kevin I-Kai; Abdulla, Waleed
	11:20-11:40	Zero-Shot Learning for Haze Removal Using Fusion of Near-Infrared and Color Images	Kato, Onhi*; Kubota, Akira
	11:40-12:00	Color Enhancement for the Colorblind Using Color Correction Intensity Map and Pix2pix Image Conversion	Komatsu, Shu*; Kubota, Akira

Session		Room	Chair
Multimedia Processing Systems in the AI Era		Meeting Room 9	-
Date	Time	Title	Authors
06-12-2024	10:40-11:00	Detecting Abnormal Machine Sounds Using An Ensemble Approach with Data Augmentation Techniques	Chan, Po-Cheng*; Lu, Chung-li; Wang, Jia-Ching
	11:00-11:20	Leveraging Semi-Supervised Learning with BEATs Feature Extraction and Bi-GRU Classification on Heterogeneous Datasets	Chen, Wei-Yu; Lu, Chung-li; Chan, Po-Cheng*; Chuang, Hsiang Feng; cheng, yu-han; Wang, Jia-Ching

	11:20-11:40	Leveraging Attention Mechanisms for Breast Cancer Diagnosis	akumalla, Brahma reddy*; Pham, Tung Bach; Zhuang, Yung-Yu; Prihasto, Bima; Chang, Pao-Chi; Wang, Jia-Ching
	11:40-12:00	Enhanced Detection of Illegally Parked Vehicles Using YOLO and Good Feature to Track Methods	Maftuh Alwafi, Fauzan; Mugi Pratama, Bobby; Le, Phuong Thi; Prihasto, Bima*; Wang, Jia-Ching

Session		Room	Chair
Few-shot Vision, Language, and Multimedia Processing under LLMs		Meeting Room 10	-
Date	Time	Title	Authors
06-12-2024	10:40-11:00	A Noisy Context Optimization Approach for Chinese Spelling Correction	Zhang, Guangwei; Xiong, Yongping; Li, Ruifan*
	11:00-11:20	GVDIE: A Zero-Shot Generative Information Extraction Method for Visual Documents Based on Large Language Models	Qi, Siyang*; Wang, Fei; Sun, Hongzhi; Ge, Yang; Xiao, Bo
	11:20-11:40	META: Text Detoxification by leveraging METAmorphic Relations and Deep Learning Methods	Choo, Alika*; Pal, Arghya; Rajanala, Sailaja; Sen, Arkendu
	11:40-12:00	Visual semantic alignment network based on pre-trained ViT for few-shot image classification	Zhang, Jiaming; Wu, Jijie; Li, Xiaoxu*