

2025 IEEE 8th International Conference on Multimedia Information Processing and Retrieval (MIPR 2025)

**San Jose, California, USA
6-8 August 2025**



**IEEE Catalog Number: CFP25K85-POD
ISBN: 979-8-3315-9466-4**

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

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

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

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

IEEE Catalog Number:	CFP25K85-POD
ISBN (Print-On-Demand):	979-8-3315-9466-4
ISBN (Online):	979-8-3315-9465-7
ISSN:	2770-4327

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

2025 IEEE 8th International Conference on Multimedia Information Processing and Retrieval (MIPR) **MIPR 2025**

Table of Contents

Message from the Organization Committee	xvii
Organizing Committee	xxi
Program Committee	xxiii

Session 1: Multimedia Data Collection, Modeling, Indexing, Storage, and Restoration

Anti-Semitic Content On Social Media Analyses Using A Hybrid Model	1
<i>Preethi Nanjundan (Christ University Pune, Lavasa, India), Ruchira Deokar (Christ University Pune, Lavasa, India), and Jayabrabu Ramakrishnan (Jazan University, Saudi Arabia)</i>	
Finding the Reflection Point: Unpadding Images to Remove Data Augmentation Artifacts in Large Open Source Image Datasets for Machine Learning	7
<i>Lucas Choi (University of Michigan, USA) and Ross Greer (University of California Merced, USA)</i>	
OpenRR-5k: A Large-Scale Benchmark for Reflection Removal in the Wild	14
<i>Jie Cai (OPPO AI Center, USA), Kangning Yang (OPPO AI Center, USA), Ling Ouyang (OPPO AI Center, USA), Lan Fu (OPPO AI Center, USA), Jiaming Ding (OPPO AI Center, USA), Jinglin Shen (OPPO AI Center, USA), and Zibo Meng (OPPO AI Center, USA)</i>	
Survey on Single-Image Reflection Removal using Deep Learning Techniques	20
<i>Kangning Yang (OPPO AI Center, USA), Huiming Sun (OPPO AI Center, USA), Jie Cai (OPPO AI Center, USA), Lan Fu (OPPO AI Center, USA), Jiaming Ding (OPPO AI Center, USA), Jinlong Li (OPPO AI Center, USA), and Zibo Meng (OPPO AI Center, USA)</i>	
EEG-Based Brain-Computer Interface for Robotic System Control Using Right-Left Hand Movements with Ensemble Model	27
<i>Yogini Dilip Borole (Marathwada Mitramandal's Institute of Technology, India), Yagnesh Challagundla (University of Florida, USA), Vinothkumar Kolluru (Stevens Institute of Technology, USA), Sachi Nandan Mohanty (VIT-AP University, India), Ming Yang (Kennesaw State University, USA), and Advaita Naidu Chintakunta (University of North Carolina, USA)</i>	

An Empirical Evaluation of Deep Learning Techniques for Clinical Video Retrieval	33
<i>Zhandi Liu (University of Missouri, USA), Mirna Becevic (University of Missouri, USA), Amy Braddock (University of Missouri, USA), Mihail Popescu (University of Missouri, USA), Eduardo J. Simoes (University of Missouri, USA), and Praveen Rao (University of Missouri, USA)</i>	

Session 2: Multimedia Retrieval

A Bird Song RAG System using Multimodal Vector Search	41
<i>Yuka Teramoto (Doshisha University, Japan) and Ryosuke Kojima (Kyoto University, Japan)</i>	
GUSL-Dehaze: A Green U-Shaped Learning Approach to Image Dehazing	45
<i>Mahtab Movahhedrad (University of Southern California, USA), Laurence Palmer (University of Southern California, USA), and C.-C. Jay Kuo (University of Southern California, USA)</i>	
Balancing Semantic Relevance and Engagement in Related Video Recommendations	52
<i>Amit Jaspal (Meta Platforms, Inc, USA), Feng Zhang (Meta Platforms, Inc, USA), Wei Chang (Meta Platforms, Inc, USA), Sumit Kumar (Meta Platforms, Inc, USA), Yubo Wang (Meta Platforms, Inc, USA), Roni Mittleman (Meta Platforms, Inc, USA), Qifan Wang (Meta Platforms, Inc, USA), and Weize Mao (Meta Platforms, Inc, USA)</i>	
Person-Centered Caption-Based Video Scene Retrieval	56
<i>Kengo Ishida (NEC Corporation, Japan), Satoshi Yamazaki (NEC Corporation, Japan), Xiaotong Ji (NEC Corporation, Japan), Junichi Funada (NEC Corporation, Japan), and Jianquan Liu (NEC Corporation, Japan)</i>	
NexusIndex: A Self-Optimizing Multimodal Framework for Fake News Detection with Dynamic Indexing and Retrieval	62
<i>Solmaz Seyed Monir (University of Washington, Tacoma), Dongfang Zhao (University of Washington, Tacoma), and Yan Bai (University of Washington, Tacoma)</i>	
Beyond General Prompts: Automated Prompt Refinement using Contrastive Class Alignment Scores for Disambiguating Objects in Vision-Language Models	69
<i>Lucas Choi (University of Michigan, USA) and Ross Greer (University of California Merced, USA)</i>	

Session 3: Media Generation with GAI

Style Transfer for Text Images with Generation Models	75
<i>Honghui Yuan (The University of Electro-Communications, Japan) and Keiji Yanai (The University of Electro-Communications, Japan)</i>	
Style Ambiguity Loss Using CLIP	82
<i>James Baker (University of Maryland, USA)</i>	
Deep Composer: Improving the String Quartet Music Generation Task	86
<i>Jacob Edward Galajda (University of Central Florida, USA), Van Hoang Bao Bach (University of Central Florida, USA), and Kien Hua (University of Central Florida, USA)</i>	

Benchmarking LLMs for Trustworthy Multimedia Retrieval in Computational Biology using Structured Zotero Graphs	93
<i>Sarvesh Tiku (Georgia Institute of Technology, USA)</i>	
FashionLIP: Automated Fashion Image Captioning Using Cascaded Contrastive Learning	97
<i>Akarsha Hegde (Yeshiva University, USA), Yudhishna Kuppala (Yeshiva University, USA), and Ming Ma (Yeshiva University, USA)</i>	
Emotion-Aware Speech Synthesis using Multimodal Deep Learning with Visual and Textual Cues	104
<i>Ketan Totlani (University of Southern California, USA), Smital Patil (KIT's College of Engineering, India), Abhijai Sasikumar (Vellore Institute of Technology, India), Fernando Moreira (REMIT, IJP, Universidade Portucalense, Porto & IEETA, Universidade de Aveiro, Aveiro; Rua Dr. Ant'onio Bernardino de Almeida, Portugal), and Sachi Nandan Mohanty (VIT-AP University)</i>	

Session 4: Machine Learning, Deep Learning, and Data Mining

IntelliCardiac: An Intelligent Platform for Cardiac Image Segmentation and Classification	109
<i>Ting Yu Tsai (University at Albany, State University of New York), An Yu (University at Albany, State University of New York), Meghana Spurthi Maadugundu (University at Albany, State University of New York), Ishrat Jahan Mohima (University at Albany, State University of New York), Umme Habiba Barsha (University at Albany, State University of New York), Mei-Hwa F. Chen (University at Albany, State University of New York), Balakrishnan Prabhakaran (University at Albany, State University of New York), and Ming-Ching Chang (University at Albany, State University of New York)</i>	
FixCLR: Negative-Class Contrastive Learning for Semi-Supervised Domain Generalization	116
<i>Ha Min Son (University of California, Davis, USA), Shahbaz Rezaei (University of California, Davis, USA), and Xin Liu (University of California, Davis, USA)</i>	
Learning Long Contexts and Forecasting with Kolmogorov-Arnold Hawkes Processes	123
<i>Naveen Kumar Pallekonda (Indian Institute of Science, India), Prabhas Reddy Onteru (Indian Institute of Science, India), Anuj Kumar (Indian Institute of Science, India), and Ambedkar Dukkupati (Indian Institute of Science, India)</i>	
QuantTune: Optimizing Model Quantization with Adaptive Outlier-Driven Fine Tuning	130
<i>Jiun-Man Chen (National Cheng Kung University, Taiwan), Yu-Hsuan Chao (National Cheng Kung University, Taiwan), Yu-Jie Wang (National Cheng Kung University, Taiwan), Ming-Der Shieh (National Cheng Kung University, Taiwan), Chih-Chung Hsu (National Yang Ming Chiao Tung University, Taiwan), and Wei-Fen Lin (Rivos Inc., USA)</i>	
Neural Network Structural Pruning and Acceleration in Frequency Domain	137
<i>Ningbo Zhu (University of Alberta, Canada), Xinyao Sun (University of Alberta, Canada), and Irene Cheng (University of Alberta, Canada)</i>	
Addressing Imbalanced Data in Stance Detection for Improved Fake News Detection	144
<i>Ruei-Hau Hsu (National Sun Yat-sen University, Taiwan), Ting-Wei Hsu (National Sun Yat-sen University, Taiwan), and Yu-Yen Chen (National Sun Yat-sen University, Taiwan)</i>	

Session 5: Deep Learning in Multimedia Data and Multimodal Fusion

Multimodal Deep Learning for Diabetic Retinopathy: A Survey	151
<i>Genyan Qin (Changde First Hospital of Traditional Chinese Medicine, China), Ning Xu (Independent Researcher, USA), and Jian Xu (University School of Medicine, China)</i>	
Text-Guided Image Invariant Feature Learning for Robust Image Watermarking	158
<i>Muhammad Ahtesham (University of Nebraska Omaha, USA) and Xin Zhong (University of Nebraska Omaha, USA)</i>	
Improving Video Moment Retrieval via LLM Augmented Nested Adapter	165
<i>Arkaprabha Bhandari (National University of Singapore, Singapore), Yongkang Wong (National University of Singapore, Singapore), Kajal Kansal (BITS Pilani Dubai Campus, UAE), Jianquan Liu (NEC Corporation, Japan), and Mohan Kankanhalli (National University of Singapore, Singapore)</i>	
Spatio-Temporal Film Grain Removal in Video Using Attention-Guided Neural Networks	172
<i>Sri Harsha Musunuri (Dolby Laboratories, USA), Anustup Choudhury (Dolby Laboratories, USA), and Guan-Ming Su (Dolby Laboratories, USA)</i>	
Modeling Event Camera Frame Sequence using Neural Field	179
<i>Anustup Choudhury (Dolby Laboratories Inc., USA) and Guan-Ming Su (Dolby Laboratories Inc., USA)</i>	
Structural MRI Synthesis for Alzheimer’s Disease via Conditional Diffusion on Anatomical Masks	186
<i>Muge Zhang (Fairleigh Dickinson University, Canada), Muhammad Ali Khaliq (University of Colorado, USA), Jamal Alsakran (Fairleigh Dickinson University, Canada), Byeong Kil Lee (University of Colorado, USA), and Jeeho Ryoo (Fairleigh Dickinson University, Canada)</i>	

Session 6: Information Coding for Content Delivery

Rate-Accuracy Bounds in Visual Coding for Machines	193
<i>Ioan V. Bajic (Simon Fraser University, Canada)</i>	
Viewport-Adaptive Scalable Tiling Multicast and Client-Side Neural Enhancement Processing of Live 360-Degree Video	200
<i>Jacob Chakareski (New Jersey Institute of Technology) and Morteza Hashemi (University of Kansas)</i>	
KaFeS: Kurtosis-Adjusted Feature Scaling for Object Tracking Tasks	207
<i>Juyeon Seo (Hanyang University, South Korea), Howon Jang (Hanyang University, South Korea), Younhee Kim (ETRI, South Korea), Jooyoung Lee (ETRI, South Korea), Se Yoon Jeong (ETRI, South Korea), and Hyunsuk Ko (Hanyang University, South Korea)</i>	

Session 7: Multimedia Systems and Infrastructures

Peer-to-Peer Synchronization for Collaborative Media Streaming	213
<i>Prathamesh Bhalange (University of Washington Bothell, USA), Siddhant Thakur (University of Washington Bothell, USA), and Min Chen (University of Washington Bothell, USA)</i>	
Hardware-Friendly Static Quantization Method for Video Diffusion Transformers	219
<i>Sanghyun Yi (California Institute of Technology, CA), Qingfeng Liu (Samsung Semiconductor Inc., CA), and Mostafa El-Khamy (Samsung Semiconductor Inc., CA)</i>	
360LiveCast: A Low-Latency and Bandwidth-Efficient Multicast Framework for Live 360 Video	226
<i>Lingzhi Zhao (University of Illinois Urbana-Champaign, USA), Qian Zhou (City University of Hong Kong, CN), Bo Chen (University of Illinois Urbana-Champaign, USA), and Klara Nahrstedt (University of Illinois Urbana-Champaign, USA)</i>	
ROI-Packing: Efficient Region-Based Compression for Machine Vision	233
<i>Eimran Hossain Eimon (Florida Atlantic University, USA), Alena Krause (Florida Atlantic University, USA), Ashan Perera (Florida Atlantic University, USA), Juan Merlos (Florida Atlantic University, USA), Hari Kalva (Florida Atlantic University, USA), Velibor Adzic (Florida Atlantic University, USA), and Borko Furht (Florida Atlantic University, USA)</i>	
Content Adaptive Multi-Scale Feature Layer Filtering	239
<i>Juan Merlos (Florida Atlantic University, USA), Eimran Eimon (Florida Atlantic University, USA), Ashan Perera (Florida Atlantic University, USA), Hari Kalva (Florida Atlantic University, USA), Velibor Adzic (Florida Atlantic University, USA), and Borko Furht (Florida Atlantic University, USA)</i>	

Session 8: Large Multimodal Models and Agents

Uncovering Latent Bias in CLIP Representations through Propensity Score Matching	243
<i>Yi-fan Hsu (Texas A&M University-Kingsville, USA)</i>	
News clip Composition with Mixture-of-Experts	250
<i>Dennis Quandt (RheinMain University of Applied Sciences, Germany), Wolfgang Ruppel (RheinMain University of Applied Sciences, Germany), and Matthias Narroschke (RheinMain University of Applied Sciences, Germany)</i>	
Multimodal Large Language Models: Developments and Directions	257
<i>Muhaimin Bin Munir (The University of Texas at Dallas, USA), Latifur Khan (The University of Texas at Dallas, USA), and Bhavani Thuraisingham (The University of Texas at Dallas, USA)</i>	
Enhancing Few-Shot Image Captioning with Discrete Region-Token Alignment and Retrieval Augmented Learning	264
<i>Koushik Nandiraju (Arksun Data Labs, USA), Sathwika Bavikadi (Rochester Institute of Technology, USA), and Alexander Loui (Rochester Institute of Technology, USA)</i>	

Degradation-Aware Image Enhancement via Vision-Language Classification	270
<i>Jie Cai (OPPO AI Center, USA), Kangning Yang (OPPO AI Center, USA), Jiaming Ding (OPPO AI Center, USA), Lan Fu (OPPO AI Center, USA), Ling Ouyang (OPPO AI Center, USA), Jiang Li (OPPO AI Center, USA), Jinglin Shen (OPPO AI Center, USA), and Zibo Meng (OPPO AI Center, USA)</i>	
A Discriminant Correlation Neural Network for Feature Representation Learning	277
<i>Lei Gao (Toronto Metropolitan University, Canada), Kai Liu (Toronto Metropolitan University, Canada), Chakkrit Termritthikun (Naresuan University, Thailand), Paisarn Muneesawang (Mahidol University, Thailand), Narit Hnoohom (Mahidol University, Thailand), and Ling Guan (Toronto Metropolitan University, Canada)</i>	

Session 9: Trustworthy AI

AI-Driven Grid Management: Deep Reinforcement Learning for Mitigating Reverse Power Flow in Distributed Photovoltaic Networks	283
<i>Xuanchen Xiang (Virginia State University, USA), Wei-bang Chen (Virginia State University, USA), Ruisheng Diao (ZJU-UIUC Institute Zhejiang University, China), Simon Foo (Florida State University, USA), and Wangjie Xu (ZJU-UIUC Institute Zhejiang University, China)</i>	
Tackling Sequential Entanglement in Split Unlearning	290
<i>Ashley Etheridge (Montclair State University), Michelle Zhu (Montclair State University), Xin Liu (Montclair State University), and Chao Huang (Montclair State University)</i>	
Mitigating Image Captioning Hallucinations in Vision-Language Models	297
<i>Fei Zhao (The University of Alabama at Birmingham, USA), Chengcui Zhang (The University of Alabama at Birmingham, USA), Runlin Zhang (University of Waterloo, Canada), Tianyang Wang (The University of Alabama at Birmingham, USA), and Xi Li (The University of Alabama at Birmingham, USA)</i>	
Progressive Coding for Privacy-Preserving Transmission on Wearable Devices	303
<i>Yuning Huang (Purdue University, USA), Long Li (University of Alabama, USA), Heather Eicher-Miller (Purdue University, USA), J. Graham Thomas (Brown University, USA), Sazonov Edward (University of Alabama, USA), and Fengqing Zhu (Purdue University, USA)</i>	
PUF-Augmented QR: Harnessing QR Code Background for Surface Authentication via Physically Unclonable Features (PUFs)	310
<i>Ethan Cornett (NC State University, USA), Will Raulston (NC State University, USA), Erik Seuster (NC State University, USA), Henry Stiff (NC State University, USA), Prasun Datta (NC State University, USA), and Chau-Wai Wong (NC State University, USA)</i>	
Non-Convex TV Regularization for Text Image Denoising with Chebyshev-Accelerated ADMM ...	316
<i>Narendra Kumar (Indian Institute of Technology, India) and Gaurav Bhatnagar (Indian Institute of Technology, India)</i>	

Session 10: Internet of Multimedia Things and User Experience

MachineStethoscope: A Smart and Cost-Effective Machine Health Monitoring System	323
<i>Beitong Tian (University of Illinois Urbana-Champaign, USA), Yaohui Wang (University of Illinois Urbana-Champaign, USA), Ahmadreza Eslaminia (University of Illinois Urbana-Champaign, USA), Ragini Gupta (University of Illinois Urbana-Champaign, USA), Robert Bruce Kaufman (University of Illinois Urbana-Champaign, USA), Leah Espenhahn (University of Illinois Urbana-Champaign, USA), Gianni Pezzarossi (University of Illinois Urbana-Champaign, USA), Mauro Sardela (University of Illinois Urbana-Champaign, USA), John Dallesasse (University of Illinois Urbana-Champaign, USA), and Klara Nahrstedt (University of Illinois Urbana-Champaign, USA)</i>	
YOLOT: A Recurrent YOLO Model for Robust Video-Based Automotive Object Detection	330
<i>Dylan Baxter (California Polytechnic State University, CA) and Jane Zhang (California Polytechnic State University, CA)</i>	
Hardware-Aware Pruning for Real-Time Multimedia Applications on Edge Devices: An Empirical Analysis	336
<i>Angad Ripudaman Singh Bajwa (HP Inc., USA) and Raghavendra B. Krishnamurthy (HP Inc., USA)</i>	
WGAN-FPN: A Weather-Adaptive GAN-Based Feature Pyramid Network for Small Aerial Object Detection in Adverse Weather Conditions	343
<i>Usman Ahmad (Zhengzhou University, China), Jing Liang (Zhengzhou University, China), Tianlei Ma (Zhengzhou University, China), Syed Waleed (Sungrow Power Supply Co., Ltd, China), Farhad Banoori (South China University of Technology, China), and Faisal Mehmood (Shenzhen University, China)</i>	
Mesh-Based Spatial Reshaping for 360° Video Compression	350
<i>Dae Yeol Lee (Dolby Laboratories, Inc., USA) and Guan-Ming Su (Dolby Laboratories, Inc., USA)</i>	
CKAN-SER: A Deep Learning Framework for Speech Emotion Recognition Using Convolutional and Kolmogorov-Arnold Networks	357
<i>Frederic Rizk (University of Louisiana at Lafayette, USA) and Chee-Hung Henry Chu (University of Louisiana at Lafayette, USA)</i>	

Session 11: Health, Sports, Culture, and Education

Interpretable Pneumonia Detection in Chest X-Rays via VICReg Self-Supervised Pretraining	363
<i>Monjurul Ahsan Prodhan (University at Albany, State University of New York, USA) and Nour Hossain (University at Albany, State University of New York, USA)</i>	
TA2N: Fast and Accurate Temporal Attention 2D Network for Remote Photoplethysmography	370
<i>Kai-Chun Huang (Feng Chia University, Taiwan), Hwai-Jung Hsu (Feng Chia University, Taiwan), and Lung Hung Chen (National Taiwan Sport University, Taiwan)</i>	

Statistical Analysis and Visualization of Medical Instruction Patterns in Multimedia-Rich Electronic Medical Records	376
<i>Miwa Sugitani (Ochanomizu University, Japan), Ryosuke Matsuo (Ochanomizu University, Japan), Tomoyoshi Yamazaki (Ochanomizu University, Japan), Kenji Araki (Ochanomizu University, Japan), Masato Oguchi (Ochanomizu University, Japan), Haruo Yokota (Josai University, Japan), and Hieu Hanh Le (Ochanomizu University, Japan)</i>	
Machine Learning-Based Sex Estimation From Human Long Bone Diaphysis Fragments Using CT Scan Data	383
<i>Talha Ibn Mahmud (Purdue University, West Lafayette, Indiana, USA), Melanie Beasley (Purdue University, West Lafayette, Indiana, USA), and Fengqing Zhu (Purdue University, West Lafayette, Indiana, USA)</i>	
Automatic Question Generation for Intuitive Learning Utilizing Causal Graph Guided Chain of Thought Reasoning	390
<i>Nicholas X. Wang (Stellar Learning Technologies, USA), Neel V. Parpia (Stellar Learning Technologies, USA), Aaryan D. Parikh (Stellar Learning Technologies, USA), and Aggelos K. Katsaggelos (Northwestern University, USA)</i>	
Multimodal Network with Instruction-Aware Memory Module for Classroom Learning Activity Recognition	395
<i>Kaiyi Li (Tsinghua University, China), Nan He (Tsinghua University, China), and Lifeng Sun (Tsinghua University, China)</i>	

Session 12: 3D Modeling, Reconstruction, and Interactive Applications

RelightableStudio: Volumetric Video Relighting using Temporal Cues and Learnable Spotlights	402
<i>Jing Yang (University of Southern California, USA), Anshul Rai (Dolby Laboratories, USA), Sandeep Bangalore Venkatesh (Dolby Laboratories, USA), Subhadra Gopalakrishnan (Dolby Laboratories, USA), and Guan-Ming Su (Dolby Laboratories, USA)</i>	
PDF-MPI: from Monocular Video to Volumetric Video	409
<i>Lingdong Wang (University of Massachusetts Amherst, USA), Dae Yeol Lee (Dolby Laboratories, USA), Guan-Ming Su (Dolby Laboratories, USA), Mohammad Hajiesmaili (University of Massachusetts Amherst, USA), and Ramesh Sitaraman (University of Massachusetts Amherst, USA)</i>	
UniHPR: Unified Human Pose Representation via Singular Value Contrastive Learning	416
<i>Zhongyu Jiang (University of Washington), Wenhao Chai (University of Washington), Lei Li (University of Copenhagen), Zhuoran Zhou (University of Washington), Cheng-Yen Yang (University of Washington), and Jenq-Neng Hwang (University of Washington)</i>	
From Pixels to Presence: High-Quality Rendering of 3D Human Models During Mobility	423
<i>Sirisha Talapuru (University of North Texas, USA), Ram Dantu (University of North Texas, USA), Vinh Quach (University of North Texas, USA), Shakila Zaman (University of North Texas, USA), and Apurba Pokharel (University of North Texas, USA)</i>	

Session 13: Novel Dataset and Feature Selection

Learning-Based Compressed-Domain Texture Recognition	430
<i>Ruolei Ji (Arizona State University, USA) and Lina J. Karam (Arizona State University, USA)</i>	
A Generic Domain Adaptation Framework Based on Kernel Inception Distance	437
<i>Igor Shtau (Navsurfvarcen Dah, USA), Daniel Suma (Navsurfvarcen Dah, USA), Ju Wang (Virginia State University, USA), Chris Payne (Virginia State University, USA), and Collin Goodloe (Virginia State University, USA)</i>	
HOIverse: A Synthetic Scene Graph Dataset With Human Object Interactions	442
<i>Mrunmai Vivek Phatak (Universität Augsburg, Germany), Julian Lorenz (Universität Augsburg, Germany), Nico Hörmann (Universität Augsburg, Germany), Jörg Hähner (Universität Augsburg, Germany), and Rainer Lienhart (Universität Augsburg, Germany)</i>	
RAISE: Realness Assessment for Image Synthesis and Evaluation	449
<i>Aniruddha Mukherjee (Kalinga Institute of Industrial Technology, India), Spriha Dubey (Indian Institute of Technology Kharagpur, India), and Somdyuti Paul (Indian Institute of Technology Kharagpur, India)</i>	

Session 14: Generative and Foundation Models

MedChat: A Multi-Agent Framework for Multimodal Diagnosis with Large Language Models	456
<i>Philip R. Liu (Purdue University, USA), Sparsh Bansal (Purdue University, USA), Jimmy Dinh (Purdue University, USA), Aditya Pawar (Purdue University, USA), Ramani Satishkumar (Purdue University, USA), Shail Desai (Purdue University, USA), Neeraj Gupta (Purdue University, USA), Xin Wang (University at Albany, State University of New York, USA), and Shu Hu (Purdue University, USA)</i>	
Robust Fairness Vision-Language Learning for Medical Image Analysis	463
<i>Sparsh Bansal (Purdue University, USA), Mingyang Wu (Purdue University, USA), Xin Wang (University at Albany - State University of New York, USA), and Shu Hu (Purdue University, USA)</i>	
Robust AI-Generated Face Detection with Imbalanced Data	470
<i>YaminiSri Krubha (Purdue University, USA), Aryana Hou (Clarkstown High School South, USA), Braden Vester (Purdue University, USA), Web Walker (Purdue University, USA), Xin Wang (University at Albany, State University of New York, USA), Li Lin (Purdue University, USA), and Shu Hu (Purdue University, USA)</i>	
To 3D or Not to 3D? Examining 3D Data Efficiency in Cross View Synthesis	477
<i>Divya Kothandaraman (University of Maryland College Park, USA), Tianyi Zhou (University of Maryland College Park, USA), Ming Lin (University of Maryland College Park, USA), and Dinesh Manocha (University of Maryland College Park, USA)</i>	

Session 15: Data Management

PROTECT: Privacy-Preserving Digital Document Forgery Detection	484
<i>Devi Listiyani (The University of Queensland, Australia), Priyanka Singh (The University of Queensland), and Xue Li (The University of Queensland, Australia)</i>	
Detecting Source Images In AI-Generated Contents Via Perceptual Hashing	491
<i>Nicholas Chenevey (California Polytechnic State University, USA) and Jane Zhang (California Polytechnic State University, USA)</i>	
A Holistic Analysis of Security and Privacy Implications of Firefox Browser Extensions	497
<i>Dyah Puji Utami (The University of Queensland, Australia), Devi Listiyani (The University of Queensland, Australia), and Priyanka Singh (The University of Queensland, Australia)</i>	
MAIA: A Taxonomy For AI-Generated Attacks	501
<i>Nitheilan Kumaar Thanigachalam (The University of Queensland, Australia), Ki Heun Kang (The University of Queensland, Australia), Priyanka Singh (The University of Queensland, Australia), and Jonathan Oliver (Broadcom Inc, Australia)</i>	
TraceNet: Segment One Thing Efficiently	508
<i>Mingyuan Wu (University of Illinois Urbana Champaign, USA), Zichuan Liu (Independent Researcher, USA), Haozhen Zheng (University of Illinois Urbana Champaign, USA), Hongpeng Guo (University of Illinois Urbana Champaign, USA), Bo Chen (University of Illinois Urbana Champaign, USA), Xin Lu (Independent Researcher, USA), and Klara Nahrstedt (University of Illinois Urbana Champaign, USA)</i>	

Session 16: Detection & Tracking

Leveraging Context from Large Multimodal Models for Weakly Supervised Video Moment Retrieval	515
<i>Sayak Nag (University of California Riverside, USA), Mahmudul Hasan (Comcast AI Comcast, USA), and Amit K Roy Chowdhury (University of California Riverside, USA)</i>	
Measuring Depth Perception for Color, Luminance, and Distance in Mixed Reality	523
<i>Hung-Jui Guo (Computer Science, National Cheng Kung University, Taiwan), Laura R. Marusich (Army Research Laboratory, U.S. DEVCOM, USA), Jonathan Z. Bakdash (Computer Science, The University of Texas at Dallas, USA), Shulan Lu (Psychology and Special Education, East Texas A&M University, USA), Andrew M. Tague (Psychology and Special Education, East Texas A&M University, USA), Reynolds J. Ballotti (Psychology and Special Education, East Texas A&M University, USA), Derek Harter (Computer Science, East Texas A&M University, USA), Omeed Eshaghi Ashtiani (Computer Science, The University of Texas at Dallas, USA), Yung-Jen Lin Guo (Department of Mathematics, National Taiwan Normal University, Taiwan), and Balakrishnan Prabhakaran (AI Plus, University at Albany, USA)</i>	

Emissivity-Guided Thermal Image Encoding: A Robust Multitask Learning Approach for Object Detection	530
<i>Fan Yang (University of Alberta, Canada), Haoyu Qiu (University of Alberta, Canada), and Irene Cheng (University of Alberta, Canada)</i>	
Visual Content Detection in Educational Videos with Transfer Learning and Dataset Enrichment	537
<i>Dipayan Biswas (University of Houston, USA), Shishir Shah (University of Houston, USA), and Jaspal Subhlok (University of Houston, USA)</i>	
Scalable Adaptive Sparse Attention Pruning: A Dynamic Approach for Efficient Neural Attention Mechanism	541
<i>Ghufron Wahyu Kurniawan (National Yang Ming Chiao Tung University, Taiwan), Jun-Wei Hsieh (National Yang Ming Chiao Tung University, Taiwan), and Chi-Chia Sun (National Taipei University, Taiwan)</i>	

Demo Session

A Web Demo Interface for Explainable Image Aesthetic Evaluation Using Vision-Language Models	547
<i>Supatta Viriyavisuthisakul (King Mongkut's University of Technology Thonburi, Thailand), Shun Yoshida (The University of Tokyo, Japan), Parinya Sanguansat (Panyapiwat Institute of Management, Thailand), and Toshihiko Yamasaki (The University of Tokyo, Japan)</i>	
VoluCapture: Multi-View Synthetic Data Capture in Unreal Engine	551
<i>Sandeep Bangalore Venkatesh (Dolby Laboratories, USA) and Guan-Ming Su (Dolby Laboratories, USA)</i>	
A Real-Time Video Quality Monitoring System for Networked Cameras	555
<i>Zhuobin Yuan (University of Cincinnati, USA) and Rui Dai (University of Cincinnati, USA)</i>	

Workshop Session: AI-SIPM

Development of an Egg Crack Detection System Using Computer Vision to Support Agricultural Development Through Modern Technology	559
<i>Nawin Kongrugsu (Muban Chombueng Rajabhat University, Thailand), Narumol Chumuang (Muban Chombueng Rajabhat University, Thailand), Todsakorn Chantongpai (Muban Chombueng Rajabhat University, Thailand), Suchat Rakchatying (Muban Chombueng Rajabhat University, Thailand), Shridhar Allagi (Kle Institute of Technology, India), Mahasak Ketcham (King Mongkut's University of Technology North Bangkok, Thailand), and Thittaporn Ganokratanaa (King Mongkut's University of Technology Thonburi, Thailand)</i>	
Interpretable Aesthetic Assessment and Prompt-Guided Image Retouching	566
<i>Supatta Viriyavisuthisakul (King Mongkut's University of Technology Thonburi, Thailand), Parinya Sanguansat (Faculty of Engineering and Technology Panyapiwat Institute of Management, Thailand), and Toshihiko Yamasaki (The University of Tokyo, Japan)</i>	

Artificial Intelligence-Driven ECG Signal Classification for Heart Disease Detection Using Decision Tree Algorithm	572
<i>Thanakrit Janchidfah (Muban Chombueng Rajabhat University, Thailand), Narumol Chumuang (Muban Chombueng Rajabhat University, Thailand), Mahasak Ketcham (King Mongkut's University of Technology North Bangkok, Thailand), Thittaporn Ganokratanaa (King Mongkut's University of Technology Thonburi Bangkok, Thailand), and Shridhar Allagi (Kle Institute of Technology, Hubil Karnataka, India)</i>	

Workshop Session: Health-MM

AI for Healing Minds: Ethical Implications of Deepfake Placebos for Emotional Support and Treatment	579
<i>Deepa Shukla (Indian Institute of Technology Jodhpur, India), Nisha Daga (Indian Institute of Technology Jodhpur, India), and George Kodimattam Joseph (Indian Institute of Technology Jodhpur, India)</i>	

Author Index	585
--------------------	-----