

2025 International Symposium on Multimedia (ISM 2025)

**Naples, Italy
8-10 December 2025**

IEEE Catalog Number: CFP25197-POD
ISBN: 979-8-3315-4591-8



Copyright © 2025, IEEE

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:	CFP25197-POD
ISBN (Print-On-Demand):	979-8-3315-4591-8
ISBN (Online):	979-8-3315-4590-1

Additional Copies of This Publication Are Available From:

Curran Associates, Inc
57 Morehouse Lane
Red Hook, NY 12571 USA

Phone:	(845) 758-0400
Fax:	(845) 758-2633
E-mail:	curran@proceedings.com
Web:	www.proceedings.com

TABLE OF CONTENTS

Facial Similarity-Guided Fine-Tuning for Hand Shape Correction in AI-Generated Human Images	1
<i>Yuki Ryu, Akira Kubota</i>	
High-Fidelity Semantic Video Communication with Controllable Image-To-Video Diffusion Models	2
<i>Cem Eteke, Alexander Griessel, Wolfgang Kellerer, Eckehard Steinbach</i>	
Diversity-Aware Active Learning for Object Detection Utilizing Time-of-Day Metadata	7
<i>Fumiya Higashide, Akira Kubota</i>	
LPConv: Laplacian Pyramid Convolutions for Parameter-Efficient Receptive Field Expansion	8
<i>Naoki Nishiya, Akira Kubota</i>	
Smarter Traps: Neural Network-Driven Classification of Small Mammals	9
<i>William Menz, Ralf Dittrich, Rakesh Rao Ramachandra Rao, Steve Göring, Alexander Raake</i>	
Exploring Privacy and Security Risks in LLMs: Data Leakage, Prompt Injection, and Membership Inference	10
<i>Giancarlo Sperlì</i>	
DWT Domain Precinct-Wise Scrambling for Encryption-then-Compression with JPEG XS	18
<i>Takayuki Nakachi, Park Cheolhwan, Yasuhisa Kato, Mitsuru Maruyama</i>	
Machine Learning Techniques for the Diagnosis and Monitoring of Nevi and Melanomas	22
<i>Giulia Di Flamminio, Fabio Persia, Daniela D'Auria, Ciro Esposito, Vincenzo Coppola</i>	
Fewer-Shot Self-Supervised Image Recoloring for Deutan Deficiency Based on Laplacian Pyramid	26
<i>Onhi Kato, Akira Kubota</i>	
QoE Evaluation of BPP Packet Wash Using ROI-Based Scalable Video Coding	28
<i>Mohammadreza Ghafari, Thibault Cholez, Olivier Festor</i>	
Evaluation of AR-Based Blind Spot Monitoring Across Diverse Driving Scenarios Using a VR Simulator	36
<i>Yohei Sakai, Tomokazu Ishikawa</i>	
A Forearm-Worn Haptic Device for Integrated Tactile and Kinaesthetic Feedback via Skin Stretch	44
<i>Selin Nur Özsert, Daniel Rodriguez-Guevara, Leonardo Franco, Wenxuan Wei, Eckehard Steinbach</i> <i>Domenico Prattichizzo</i>	
Prompted Vs. Organic Customer Insight: Comparing the Value of Focus Groups and Online Reviews in Product and Service Innovation	51
<i>Maren Schnieder, Ana Isabel Canhoto, Ramin Behbehani, Ahmad Beltagui, Niraj Kumar</i> <i>Amirreza Alizamani</i>	
Low-Light Image Enhancement with Adaptive Brightness Transform Models for Video See-Through AR	59
<i>Yingen Xiong, Christopher Peri</i>	
Video Classification of Marchantia Polymorpha Using a Video Vision Transformer with Emphasized Channel Information	61
<i>Haruhiko Murata, Naoki Minamino, Takashi Ueda, Yohei Kondo, Kazuhiro Hotta</i>	

Syntax-Aware Transformer for Sentiment Analysis of Japanese SNS Text	65
<i>Sotaro Shiozawa, Akira Kubota</i>	
Coding Gaussian Splat Scenes with V3C/V-PCC	69
<i>Patrice Rondao Alface, Lauri Ilola, Lukasz Kondrad</i>	
Comparative Evaluation of Deep Learning Methods for Wood Surface Defect Detection: A Comprehensive Study of Semantic Segmentation Approaches	74
<i>Yuki Yanai, Tomokazu Ishikawa</i>	
Exploiting LLMs for Metadata-Based Video Quality Prediction	78
<i>Steve Göring, Rakesh Rao Ramachandra Rao, Alexander Raake</i>	
Lightweight High-Accuracy Tomato Detection and Classification by Efficientnet-Enhanced YOLOv8	80
<i>Hayato Tsukada, Akira Kubota</i>	
Evaluation of a Floating-Head Communication Prototype for Video-Conferencing	81
<i>William Menz, Alexander Zoubarev, David Kutschke, Rakesh Rao Ramachandra Rao, Louay Bassbous Sven Bliedung von der Heide, Steve Göring, Alexander Raake</i>	
Rendering Compressed Point Clouds with a Voxel-Based Method	82
<i>HyungWoo Kang, YeoJun Yoon, Joong-Hwan Baek, Byung Tae Oh</i>	
Application of Computer Vision Research (ISVP.AI) in the Development of the Comprehensive Stellis One Platform for Sports Organizations	83
<i>Łukasz Gq̄siorowski, Jagoda Lazarek, Sebastian Purtak, Paweł Gora</i>	
Recognition of Pitching Habits Using Multimodal Data of RGB Video and Skeleton	84
<i>Satoki Hidaka, Kazuhiro Hotta</i>	
Layout-Aware Self-Correcting Prompts for Multimodal LLM Parking Lot Monitoring	88
<i>Viviana Crescitelli</i>	
Comparative Analysis of Face Recognition Models: Runtime Environments and Compute Units on Edge	92
<i>Łukasz Grzymkowski, Tomasz P. Stefański</i>	
An Agent-Driven Architecture for Harmful Meme Detection through Multimodal Decomposition	96
<i>Gian Marco Orlando, Marco Perillo, Diego Russo, Vincenzo Moscato</i>	
Attention-Enhanced Multi-Branch Spiking Neural Network for Event Stream Super-Resolution	103
<i>Ahmadreza Sezavar, Catarina Brites, João Ascenso</i>	
SynthMed: Generating and Detecting Multimodal Deepfakes for Healthcare Communication	108
<i>Mariano Barone, Francesco Di Serio, Vincenzo Moscato, Marco Postiglione, Giuseppe Riccio Antonio Romano</i>	
Secure AI-Driven Super-Resolution for Real-Time Mixed Reality Applications	116
<i>Mohammad Waqas Usmani, Sankalpa Timilsina, Michael Zink, Susmit Shannigrahi</i>	
Personalised Stress Detection: An Exploration of Temporal Multimodal Late Fusion Strategies	124
<i>Misha Libman, Gelareh Mohammadi</i>	
Extrinsic Calibration of RGB-D Cameras Using Depth Refinement	133
<i>Peter O. Fasogbon</i>	

Evaluating the Emerging MPEG Video Coding for Machines in Semantic Segmentation	141
<i>Khoa Pham-Dinh, Farhad Pakdaman, Honglei Zhang, Hamed Rezazadegan Tavakoli, Nam Le Jukka Ahonen, Moncef Gabbouj</i>	
Adaptive Obfuscation for Reusing RGB Datasets for Privacy-Preserving Human Pose Estimation	149
<i>Francesco Pistolesi, Matteo Mugnai, Beatrice Lazzerini</i>	
Dialogue-Pseudo: A Speaker Pseudonymization Framework for Privacy Protection in Dialogue Speech Data	156
<i>Aoi Ito, Katunobu Itou</i>	
Opt360: QoE Optimization for 360° Video Streaming	164
<i>Reza Hedayati, Mea Wang, Logan Rakai</i>	
Foot-Strike Pattern Recognition from Inertial Data with Machine Learning	172
<i>Michele Baldassini, Francesco Pistolesi, Beatrice Lazzerini</i>	
The Sustainability Card: Measuring Sustainability of Multimedia AI Models	180
<i>Francesco Pistolesi, Michele Baldassini, Matteo Mugnai, Beatrice Lazzerini</i>	
One Size Doesn't Fit All: Age-Aware Gamification Mechanics for Multimedia Learning Environments	182
<i>Sarah Kaißer, Markus Kleffmann, Kristina Schaaff</i>	
An Influence Analysis of Hybrid Lectures with a Simple Setup on the Student Experience	190
<i>Florian Schimanke, Robert Mertens, Felix Prankel</i>	
RAG Chatbots for Educational Virtual Field Trips	196
<i>Suryaprakash Reddy Kalvakolu, Heinrich Söbke, Florian Wehking, Mukesh Chandra Kumar Mamidala Eckhard Kraft</i>	
An Efficient Optimization Criterion for Multi-View Feature Representation Learning	202
<i>Lei Gao, Kai Liu, Kevin Tang, Ling Guan</i>	
AMICO: A Semantic and Multimodal Framework for AI-Assisted Clinical Reporting	208
<i>Antonio Laudante, Mariano Barone, Giuseppe Riccio, Antonio Romano, Francesco Di Serio Antonio Scialdone, Francesco Porciello, Nicola Rainone, Vincenzo Moscato</i>	
Empowering Access to Public Services: An Analysis on Multimodal, Retrieval-Augmented Chatbots for Indic Language Support to Farmers	214
<i>Mohsina Bilal, Gopakumar G</i>	
Analysis of Multimodal LLMs in VQA in the Field of Radiology	218
<i>Cristovão Pessoa Cândido Neto, Cláudio de Souza Baptista, André Luiz F. Alves, Vivek Swarnakar Anselmo Cardoso de Paiva</i>	
Extending Visual Dialog Beyond English: An Analysis of Monolingual and Multilingual Models	226
<i>Milena Menezes Adão, Silvio Jamil F. Guimarães, Zenilton K. G. Patrocínio</i>	
On the Suitability of Perceptual Quality Metrics for Learning-Based Screen Content Compression	234
<i>H. Burak Dogaroglu, Hongjie You, Atanas Boev, Elena Alshina, Eckehard Steinbach</i>	
Multiscale RGB–Thermal Fusion for Vulnerable Road User Detection with ScaleFuse	238
<i>İbrahim Tinas, Yavuz Selim Bostancı, Mujdat Soyuturk</i>	

Graph-Based Evaluation of Visual Brain Decoding from fMRI Data	244
<i>Mohammad Moradi, Morteza Moradi, Marco Grassia, Giuseppe Mangioni</i>	
CCAFF: Object Tracking Under Heavy Occlusion	252
<i>Abdul Bhutta, Naimul Khan, Ling Guan</i>	
Physics-Guided Exposure Parameter Estimation for Image Metadata Verification	258
<i>Sharmilee Rajkumar Rajan, Ming-Ching Chang, Pradeep K. Atrey</i>	
Extracting Player Speed from Football Videos	266
<i>Ole Kristian Rustebakke, Mehdi Houshmand Sarkhoosh, Cise Midoglu, Pål Halvorsen</i>	
ExposureEngine: Oriented Logo Detection and Sponsor Visibility Analytics in Sports Broadcasts	274
<i>Mehdi Houshmand Sarkhoosh, Frøy Øye, Henrik Nestor Sørli, Nam Hoang Vu, Dag Johansen</i> <i>Cise Midoglu, Tomas Kupka, Pål Halvorsen</i>	
MCAD: Multimodal Context-Aware Audio Description Generation for Soccer	280
<i>Lipisha Chaudhary, Trisha Mittal, Subhadra Gopalakrishnan, Ifeoma Nwogu, Jaclyn Pytlarz</i>	
Knowledge-Based Behavioral Biometrics for Secure Authentication in Virtual Reality	288
<i>Numan Zafar, Priyo Ranjan Kundu Prosun, Shafique Ahmad Chaudhry</i>	
From 2D to 3D: How Discrete Dependencies Enable Cross-Dimensional Inference in Neural Networks in Defiance of Euclidean Geometry	292
<i>Gerald Friedland, Robert Mertens</i>	
Forecasting “Neg Storms”: Time-Aware Modeling of Toxic Situations in Social Media	300
<i>Irien Akter, Vivek K. Singh, Pradeep K. Atrey</i>	
Personalized Adaptive Magnification in Gaze-Based Interaction	308
<i>Florian Eggenkemper, Jana Swerew, Teresa Reher, Manuel Hanhoff, Constantin A. Rothkopf</i> <i>Robert Mertens</i>	
Adaptation of CDN at the Edge Using Cloud-Native Network Telemetry Across Media Scenarios	315
<i>Javier Iglesias, Felipe Mogollón, Iñigo Tamayo, Zaloa Fernández, Olov Danielsson, Ivan Pretel</i> <i>Asier Lopez</i>	
Quality Assessment of Dynamic 3D Model in Virtual Reality: Effects of Level of Detail and Viewing Distance	323
<i>Duc V. Nguyen, Nguyen Thi Quynh Ly, Truong Thu Huong</i>	
3GPP PDU Set Framework: Release 19 Updates	327
<i>Serhan Gül, Igor D.D. Curcio</i>	
A One-Class Structural Similarity-Based Autoencoder for the Detection of Malaria-Infected Cells	331
<i>Moses Omondi, Yassine Belkhouche</i>	
Comparison of Multimodal Fall Detection Strategies	336
<i>Reema Maheshbhai Gadhia, Nasim Hajari</i>	
Emotion Detection and Classification of Different Saudi Dialects	340
<i>Rehab K Qarout, Joud Y Samkari, Rahaf M ALFudhayl, Ruba H ALSulami, Shada M Basudan</i> <i>Ghadi K AlJuhani, Nuha Zamzami</i>	

AR in HbbTV-Based Hybrid TV Services	346
<i>Fernando Boronat, Lluc Simó, Rubén Prieto, Almanzor Sapena</i>	
A Hybrid Noise Perturbation-Based Denoising Autoencoder for Machine Sound Anomaly Detection	354
<i>Kadir Torun, Mustafa Sert</i>	
A First Look at Open-GoP Streaming with Av1 S-Frames	358
<i>Akram Ansari, Ali John Naqvi, Mea Wang, Emir Halepovic</i>	
On Progressive Compressed Neural Model Storage	366
<i>Hamed R. Tavakoli, Homayun Afrabadpey</i>	
Metadata-Guided Hot Swapping of Specialized Super-Resolution Models in Streaming Systems	370
<i>Alperen F. Zengin, Ekrem Cetinkaya, Ali C. Begen, Saba Ahsan, Serhan Gül, Kashyap Kammachi-Sreedhar Emre Aksu</i>	
TARS: Temporal-Spatial Adaptation for Volumetric Video Streaming	375
<i>Hadi Heidarirad, Amir Allahveran, Mea Wang</i>	
Cost-Optimal Design of Hybrid Broadcast—Unicast Video Delivery Systems	383
<i>Yuriy A. Reznik</i>	