

2025 19th International Conference on Machine Vision and Applications (MVA 2025)

**Kyoto, Japan
26-28 July 2025**



**IEEE Catalog Number: CFP25A66-POD
ISBN: 979-8-3315-1497-6**

**Copyright © 2025, IEICE
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:	CFP25A66-POD
ISBN (Print-On-Demand):	979-8-3315-1497-6
ISBN (Online):	978-4-8855-2355-7

Additional Copies of This Publication Are Available From:

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

CURRAN ASSOCIATES INC.
proceedings
.com

TABLE OF CONTENTS

FMDP: Leveraging a Foundation Model for Dual-Pixel Disparity Estimation.....	1
<i>Doehyung Lee, Zhuofeng Wu, Yusuke Monno, Masatoshi Okutomi</i>	
Capturing Fine-Grained Alignments Improves 3D Affordance Detection.....	6
<i>Junsei Tokumitsu, Yuiga Wada</i>	
Self-Supervised 3D Image Deburring for Lattice Light Sheet Microscopy	13
<i>Ziyi Chen, Shiro Suetsugu, Yuko Mimori-Kiyosue, Hajime Nagahara</i>	
ZECO: ZeroFusion Guided 3D MRI Conditional Generation.....	19
<i>Feiran Wang, Bin Duan, Jiachen Tao, Nikhil Sharma, Dawen Cai, Yan Yan</i>	
Advancing Disease Detection Using Deep Learning in Low-Data Environments	24
<i>Jian-Ming Lu, Chien-Fu Chen</i>	
DLSF: Dual-Layer Synergistic Fusion for High-Fidelity Image Synthesis.....	30
<i>Zhen-Qi Chen, Yuan-Fu Yang</i>	
Data-Driven Head Motion Generation Through Natural Gaze-Head Coordination.....	36
<i>Xiaohan Liu, Yilin Wen, Yusuke Sugano</i>	
Low-Latency Real-Time Audio-Driven Talking Head Generation Based on Future Speech Feature Prediction	41
<i>Weijie Guo, Naoshi Kaneko, Yoshiaki Akazawa</i>	
Leveraging Masked Feature and Consistency Regularization for Unsupervised Domain Adaptation Based Semi-Supervised Semantic Segmentation	46
<i>Chi-Teng Yeh, Hao-Ting Li, Kuo-En Chen, Chen-Kuo Chiang</i>	
MoExDA: Domain Adaptation for Edge-Based Action Recognition	51
<i>Takuya Sugimoto, Ning Ding, Toru Tamaki</i>	
MobileSACNet: Lightweight Spectral-Spatial Compression for Hyperspectral Segmentation in Autonomous Driving Systems.....	58
<i>Chao Wang, Daisuke Miyazaki</i>	
Supervised Domain Adaptation from Scene Text Recognition for Licence Plate Recognition	63
<i>Novanto Yudistira, Wahyu Valentino Marasitua</i>	
CLIP-Guided Cross-Modal Feature Fusion Based Few-Shot Learning for Nighttime Pavement Defect Detection.....	68
<i>Chongchong Zan, Songlin Du, Takeshi Ikenaga</i>	
Decoupled Scale and Appearance for Optimal Deep Diamond ReID	73
<i>Arno Waes, Bram Claes, Toon Goedemé</i>	
IRR-RADA: A Reflection-Aware Saliency Map and Adaptive Curriculum Learning Based Data Augmentation Method for Image Reflection Removal	79
<i>Wuyou Zhou, Fengshan Zhao, Yujie Wang, Qin Liu, Takeshi Ikenaga</i>	
Simple Yet Effective Way to Use Polarimetric Information in Stereo Matching.....	84
<i>Jinyu Zhao, Yusuke Monno, Masatoshi Okutomi</i>	

Scene Recognition Meets Knowledge Graphs: Enhancing Robustness to Object Diversity	89
<i>Keisuke Yamashita, Seiya Ito, Tomoaki Yamazaki, Kouzou Ohara</i>	
An Automatic Rating Approach Using Machine Learning and Feature Selection for Finger Tapping in MDS-UPDRS Part III.....	94
<i>Yi-Hung Chiu, Tung-Kuan Liu, Chih-Ping Yang, Ching-Fang Chien, Li-Min Liou, Lung- Chang Lin, Huei-Ping Dong, Chen-Sen Ouyang</i>	
Noise-Based Regularized Training for Diffusion Models	98
<i>Yuzhang Shang, Yubin Lu, Jinchao Feng, Ming Zhong, Yan Yan</i>	
Unsupervised 3D Braided Hair Reconstruction from a Single-View Image	104
<i>Jing Gao</i>	
Enhancing Reliability of Medical Image Diagnosis Through Top-Rank Learning with Rejection Module	110
<i>Xiaotong Ji, Ryoma Bise, Seiichi Uchida</i>	
Pre-Manipulation Alignment Prediction for Open-Vocabulary Object Manipulation Based on End- Effector Trajectories	115
<i>Motonari Kambara, Komei Sugiura</i>	
Domain Generalization of Pathological Image Segmentation by Patch-Level and WSI-Level Contrastive Learning	120
<i>Yuki Shigeyasu, Shota Harada, Akihiko Yoshizawa, Kazuhiro Terada, Naoki Nakazima, Mariyo Kurata, Hiroyuki Abe, Tetsuo Ushiku, Ryoma Bise</i>	
Dynamic Age Estimation Via Mixture of Experts: Bridging Semantic and Structural Models	125
<i>Yu-Ling Chen, Yu-Hsuan Chiu, Ching-Hsun Chang, Gee-Sern Hsu</i>	
Temporal Conditioning for Realistic Performance Video Generation from Instrumental Sounds.....	131
<i>Yusaku Nakanose, Chito Nakagawa, Katsufumi Inoue, Michifumi Yoshioka</i>	
Binned MSE for Imbalanced Dust Density Estimation	136
<i>Andreas Michel, Jannick Kuester, Jannik Sheikh, Wolfgang Gross, Martin Weinmann</i>	
IG-ODAM: Instance-Aware Visual Explanations for Object Detection with Integrated Gradients	141
<i>Yuma Nakai, Tsubasa Hirakawa, Takayoshi Yamashita, Hironobu Fujiyoshi</i>	
ShadowAug: A Multi-Strategy Data Augmentation Method for Image Shadow Removal	146
<i>Yujie Wang, Fengshan Zhao, Wuyou Zhou, Zheyi Li, Ziyang Li, Qin Liu, Takeshi Ikenaga</i>	
Any-Scale Object Detection Using Arbitrary-Scaled Images.....	151
<i>Kazutoshi Akita, Norimichi Ukita</i>	
Bidirectional Action Sequence Learning for Long-Term Action Anticipation with Large Language Models.....	157
<i>Yuji Sato, Yasunori Ishii, Takayoshi Yamashita</i>	
3D Object Reconstruction Through Integration of Hyperspectral and RGB-D Imaging.....	163
<i>Cheng Hong, Guillaume Caron, Nevrez Imamoglu, Fumio Kanehiro</i>	
Parallel Sampling of Diffusion Models on SO(3)	168
<i>Yan-Ting Chen, Hao-Wei Chen, Tsu-Ching Hsiao, Chun-Yi Lee</i>	
Style-Preserving Diffusion for Scene Text Editing.....	173
<i>Wei-Zhe Jian, Gee-Sern Hsu</i>	

FlowLoss: Dynamic Flow-Conditioned Loss Strategy for Video Diffusion Models.....	178
<i>Kuanting Wu, Kei Ota, Asako Kanezaki</i>	
RGB-Thermal Cooperative Robot Vision Strategy for Multi-Person Tracking in Both Well-Lit and Low-Light Scenes.....	184
<i>Tatsuro Sakai, Kanji Tanaka, Jonathan Tay Yu Liang, Muhammad Adil Luqman, Daiki Iwata</i>	
A Minimalist Approach to HDR Image Compression with Applications to Low-Light Image Enhancement	190
<i>Alexander Belyaev, Pierre-Alain Fayolle, Shin Yoshizawa, Hideo Yokota</i>	
Magic for the Age of Quantized DNNs	195
<i>Ryuji Saiin, Kazuma Suetake, Yoshihide Sawada</i>	
A Lightweight Convolutional Neural Network for Underwater Image Quality Enhancement.....	201
<i>Ming-Wei Chen, Chia-Yeh Lin, Shun-Pin Hsu</i>	
Real-Time Fire Detection Using Hybrid Feature Extraction: Color, Texture, and Motion Analysis	206
<i>Waffiq Maaraja, Aufaclav Zatu Kusuma Frisky, Fadillah Siva, Hanum Khairana Fatmah</i>	
Viewpoint-Aware 3D Dense Captioning	212
<i>Yuta Irisawa, Seiya Ito, Tomoaki Yamazaki, Ken Sakurada, Ryuhei Hamaguchi, Masaki Onishi, Kouzou Ohara</i>	
Transformer-Based Visual Grounding with Inter-Modality Cross Attention.....	217
<i>Yu-Hong Zheng, Guo-Shiang Lin, Ku-Yaw Chang</i>	
Age Prediction of Komatsuna Using Hu Moments with Neural Networks for Small Datasets.....	223
<i>Moeri Okuda, Shinsaku Hiura</i>	
Statistic Temporal Checking and Depth Layering Based Multi-Object Relative Size Estimation from Monocular Video	227
<i>Ziyue Wang, Xina Cheng, Takeshi Ikenaga</i>	
Edge-Augmented HLAC and Gaussian Distribution-Based Weighted Feature Extraction for 1-ms Abnormal Detection System in Logistics	232
<i>Yuanfeng Zhang, Wenshuo Wu, Tingting Hu, Manabu Inoue, Hiroshi Mitani, Takeshi Ikenaga</i>	
Geometrically Constrained Position Estimation Through Low-Level Tracking	237
<i>Henry Wood, Andrew Mills, William R Jacobs, Rodrigo Jimenez</i>	
Very Similar Appearance Feature Classification for Chronic Endometritis Diagnosis in Hysteroscopy Images.....	242
<i>Yuki Yoshida, Yuki Sawada, Takeshi Sato, Mayumi Sugiura-Ogasawara, Norimichi Ukita</i>	
Detecting Hand-Object Interaction Based on Movements in Hand Surrounding Region	247
<i>Rikuto Konishi, Taichi Uchimura, Toru Abe</i>	
Multi-Person Pose Estimation Evaluation Using Optimal Transportation and Improved Pose Matching.....	252
<i>Takato Moriki, Hiromu Taketsugu, Norimichi Ukita</i>	
Object State Recognition in Cooking Videos Through End State Frames Analysis	257
<i>Yutaro Kitakami, Katsufumi Inoue, Michifumi Yoshioka</i>	

Detection of Medial Epicondyle Avulsion in Elbow Ultrasound Images Via Bone Structure Reconstruction.....	262
<i>Shizuka Akahori, Shotaro Teruya, Pragyana Shrestha, Yuichi Yoshii, Satoshi Izuka, Akira Ikumi, Hiromitsu Tsuge, Itaru Kitahara</i>	
Point Cloud Edge Extraction Based on 3D Point Separability Filter with Spherical Mask.....	267
<i>Gulpi Qorik Oktagalu Pratamasunu, Naish Ibrahim, Hiroaki Kawamoto, Kazuhiro Fukui</i>	
Gaze Attention Estimation for Medical Environments.....	272
<i>Natchapol Shinno, Yuki Furuya, Takeshi Saitoh, Haibo Zhang, Keiko Tsuchiya, Hitoshi Sato, Frank Coffey</i>	
Revisiting Self-Generating Simple Figure Patterns for Learning Microscopy Image Segmentation	277
<i>Shuhei Iijima, Takuji Ohashi, Yu-Chieh Lin, Youyang Ng</i>	
Modality Selection and Skill Segmentation Via Cross-Modality Attention	282
<i>Jiawei Jiang, Kei Ota, Devesh K. Jha, Asako Kanezaki</i>	
Efficient Skeleton-Based Action Recognition Using Superposed Shape Subspace	288
<i>Santos Enoque Safrão, Kazuhiro Fukui</i>	
Semantic Segmentation of iPS Cells: Case Study on Model Complexity in Biomedical Imaging.....	293
<i>Maoquan Zhang, Bisser Raytchev, Xiujuan Sun</i>	
Snapshot Hyperspectral Imaging Using Petrographic Thin Section.....	298
<i>Yuta Fujimoto, Kazuya Kitano, Kazuma Fujiwara, Yuki Fujimura, Takuya Funatomi, Yasuhiro Mukaigawa</i>	
Cross-Modal Knowledge Distillation from First-Person Views to Third-Person BEV Maps for Universal Point Goal Navigation.....	303
<i>Riku Uemura, Kanji Tanaka, Daiki Iwata, Kenta Tsukahara, Daigo Nakajima, Tomoe Hiroki</i>	
Impact of Optical System Size on Robustness in Laser Speckle Authentication	308
<i>Naoki Fujieda, Kazuya Kitano, Takuya Funatomi, Rui Ishiyama, Yasuhiro Mukaigawa</i>	
Modifying Generative Distributions in Latent Diffusion Models to Improve Alignment with Desired Properties	313
<i>Sven Sickert, Maria Gogolev, Niklas Penzel, Tim Büchner, Joachim Denzler</i>	
MVA 2025 Small Multi-Object Tracking for Spotting Birds Challenge: Dataset, Methods, and Results	319
<i>Yuki Kondo, Norimichi Ukita, Riku Kanayama, Yuki Yoshida, Takayuki Yamaguchi, Xiang Yu, Guang Liang, Xinyao Liu, Guan-Zhang Wang, Wei-Ta Chu, Bing-Cheng Chuang, Jia-Hua Lee, Pin-Tseng Kuo, I-Hsuan Chu, Yi-Shein Hsiao, Cheng-Han Wu, Po-Yi Wu, Jui-Chien Tsou, Hsuan-Chi Liu, Chun-Yi Lee, Yuan-Fu Yang, Kosuke Shigematsu, Asuka Shin, Ba Tran</i>	
Intersection-Based Ensemble for Small Multi-Object Tracking in Challenging Environments	334
<i>Guan-Zhang Whang, Wei-Ta Chu</i>	
Boosting Small Object Tracking Via Collaborative Detection Transformer.....	340
<i>Bing-Cheng Chuang, Jia-Hua Lee, Pin-Tseng Kuo, I-Hsuan Chu, Yi-Shein Hsiao, Cheng-Han Wu, Po-Yi Wu, Jui-Chien Tsou, Hsuan-Chi Liu, Chun-Yi Lee, Yuan-Fu Yang</i>	
Confidence-Based Adaptive Weighted Boxes Fusion for Multi-Object Tracking of Small Birds.....	346
<i>Kosuke Shigematsu, Asuka Shin</i>	

Author Index