

39th AAAI Conference on Artificial Intelligence (AAAI-25), 37th Conference on Innovative Applications of Artificial Intelligence (IAAI-25), and 15th Symposium on Educational Advances in Artificial Intelligence (EAAI-25)

Volume 4: AAAI Technical Tracks

- Computer Vision III

Philadelphia, Pennsylvania, USA
25 February - 4 March 2025

Part 1 of 2

ISBN: 979-8-3313-2514-5

Printed from e-media with permission by:

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



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2025) by Association for the Advancement of Artificial Intelligence
All rights reserved.

Printed with permission by Curran Associates, Inc. (2025)

For permission requests, please contact Association for the Advancement of Artificial Intelligence
at the address below.

Association for the Advancement of Artificial Intelligence
2275 East Bayshore Road
Suite 160
Palo Alto, California 94303
USA

Phone: 1-650-328-3123
Fax: 1-650-321-4457

<https://aaai.org/Press/press.php>

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
Email: curran@proceedings.com
Web: www.proceedings.com

TABLE OF CONTENTS

PART 1

AAAI TECHNICAL TRACK ON COMPUTER VISION III

BUFF: Bayesian Uncertainty Guided Diffusion Probabilistic Model for Single Image Super-Resolution.....	3474
<i>Zihao He, Shengchuan Zhang, Runze Hu, Yunhang Shen, Yan Zhang</i>	
Robust and Consistent Online Video Instance Segmentation via Instance Mask Propagation.....	3483
<i>Miran Heo, Seoung Wug Oh, Seon Joo Kim, Joon-Young Lee</i>	
Generalized Class Discovery in Instance Segmentation.....	3491
<i>Cuong Manh Hoang, Yeejin Lee, Byeongkeun Kang</i>	
WildFake: A Large-Scale and Hierarchical Dataset for AI-Generated Images Detection	3500
<i>Yan Hong, Jianming Feng, Haoxing Chen, Jun Lan, Huijia Zhu, Weiqiang Wang, Jianfu Zhang</i>	
FashionTailor: Controllable Clothing Editing for Human Images with Appearance Preserving.....	3509
<i>Jie Hou, Jianghong Ma, Xiangyu Mu, Haijun Zhang, Zhao Zhang</i>	
Prompt Tuning In a Compact Attribute Space	3518
<i>Shiyu Hou, Tianfei Zhou, Shuai Zhang, Ye Yuan, Guoren Wang</i>	
ZeroMamba: Exploring Visual State Space Model for Zero-Shot Learning	3527
<i>Wenjin Hou, Dingjie Fu, Kun Li, Shiming Chen, Hehe Fan, Yi Yang</i>	
BloomScene: Lightweight Structured 3D Gaussian Splatting for Crossmodal Scene Generation	3536
<i>Xiaolu Hou, Mingcheng Li, Dingkang Yang, Jiawei Chen, Ziyun Qian, Xiao Zhao, Yue Jiang, Jinjie Wei, Qingyao Xu, Lihua Zhang</i>	
Training-and-Prompt-Free General Painterly Harmonization via Zero-Shot Disentanglement on Style and Content References.....	3545
<i>Teng-Fang Hsiao, Bo-Kai Ruan, Hong-Han Shuai</i>	
GaussianSR: High Fidelity 2D Gaussian Splatting for Arbitrary-Scale Image Super-Resolution.....	3554
<i>Jintong Hu, Bin Xia, Bin Chen, Wenming Yang, Lei Zhang</i>	
VRVVC: Variable-Rate NeRF-Based Volumetric Video Compression	3563
<i>Qiang Hu, Houqiang Zhong, Zihan Zheng, Xiaoyun Zhang, Zhengxue Cheng, Li Song, Guangtao Zhai, Yanfeng Wang</i>	
MonoBox: Tightness-Free Box-Supervised Polyp Segmentation Using Monotonicity Constraint.....	3572
<i>Qiang Hu, Zhenyu Yi, Ying Zhou, Fan Huang, Mei Liu, Qiang Li, Zhiwei Wang</i>	
Exploiting Multimodal Spatial-temporal Patterns for Video Object Tracking	3581
<i>Xiantao Hu, Ying Tai, Xu Zhao, Chen Zhao, Zhenyu Zhang, Jun Li, Bineng Zhong, Jian Yang</i>	
Motion Decoupled 3D Gaussian Splatting for Dynamic Object Representation.....	3590
<i>Xiao Hu, Libo Long, Jochen Lang</i>	

V2Xum-LLM: Cross-Modal Video Summarization with Temporal Prompt Instruction Tuning.....	3599
<i>Hang Hua, Yunlong Tang, Chenliang Xu, Jiebo Luo</i>	
Identity-Text Video Corpus Grounding	3608
<i>Bin Huang, Xin Wang, Hong Chen, Houlun Chen, Yaofei Wu, Wenwu Zhu</i>	
SubjectDrive: Scaling Generative Data in Autonomous Driving via Subject Control	3617
<i>Binyuan Huang, Yuqing Wen, Yucheng Zhao, Yaosi Hu, Yingfei Liu, Fan Jia, Weixin Mao, Tiancai Wang, Chi Zhang, Chang Wen Chen, Zhenzhong Chen, Xiangyu Zhang</i>	
HUANG: A Robust Diffusion Model-based Targeted Adversarial Attack Against Deep Hashing Retrieval	3626
<i>Chihan Huang, Xiaobo Shen</i>	
LPCG: A Self-conditional Architecture for Labeled Point Cloud Generation	3635
<i>Dongshuo Huang, Xiaoshui Huang, Chengdong Zhang, Yilei Shi</i>	
FatesGS: Fast and Accurate Sparse-View Surface Reconstruction Using Gaussian Splatting with Depth-Feature Consistency	3644
<i>Han Huang, Yuhun Wu, Chao Deng, Ge Gao, Ming Gu, Yu-Shen Liu</i>	
Densely Connected Parameter-Efficient Tuning for Referring Image Segmentation	3653
<i>Jiaqi Huang, Zunnan Xu, Ting Liu, Yong Liu, Haonan Han, Kehong Yuan, Xiu Li</i>	
Wavelet-Assisted Multi-Frequency Attention Network for Pansharpening.....	3662
<i>Jie Huang, Rui Huang, Jinghao Xu, Siran Peng, Yule Duan, Liang-Jian Deng</i>	
AUTE: Peer-Alignment and Self-Unlearning Boost Adversarial Robustness for Training Ensemble Models.....	3671
<i>Lifeng Huang, Tian Su, Chengying Gao, Ning Liu, Qiong Huang</i>	
EvoChart: A Benchmark and a Self-Training Approach Towards Real-World Chart Understanding.....	3680
<i>Muye Huang, Han Lai, Xinyu Zhang, Wenjun Wu, Jie Ma, Lingling Zhang, Jun Liu</i>	
VProChart: Answering Chart Question Through Visual Perception Alignment Agent and Programmatic Solution Reasoning	3689
<i>Muye Huang, Lingling Zhang, Han Lai, Wenjun Wu, Xinyu Zhang, Jun Liu</i>	
SLIP: Spoof-Aware One-Class Face Anti-Spoofing with Language Image Pretraining.....	3697
<i>Pei-Kai Huang, Jun-Xiong Chong, Cheng-Hsuan Chiang, Tzu-Hsien Chen, Tyng-Luh Liu, Chiou-Ting Hsu</i>	
Resolving Multi-Condition Confusion for Finetuning-Free Personalized Image Generation	3707
<i>Qihan Huang, Siming Fu, Jinlong Liu, Hao Jiang, Yipeng Yu, Jie Song</i>	
Unleashing the Temporal-Spatial Reasoning Capacity of GPT for Training-Free Audio and Language Referenced Video Object Segmentation	3715
<i>Shaofei Huang, Rui Ling, Hongyu Li, Tianrui Hui, Zongheng Tang, Xiaoming Wei, Jizhong Han, Si Liu</i>	
ZoRI: Towards Discriminative Zero-Shot Remote Sensing Instance Segmentation	3724
<i>Shiqi Huang, Shuting He, Bihan Wen</i>	
DreamPhysics: Learning Physics-Based 3D Dynamics with Video Diffusion Priors	3733
<i>Tianyu Huang, Haoze Zhang, Yihan Zeng, Zhilu Zhang, Hui Li, Wangmeng Zuo, Rynson W. H. Lau</i>	

Efficient Indoor Depth Completion Network Using Mask-adaptive Gated Convolution.....	3742
<i>Tingxuan Huang, Jiacheng Miao, Shizhuo Deng, Tong Jia, Dongyue Chen</i>	
Manta: Enhancing Mamba for Few-Shot Action Recognition of Long Sub-Sequence	3751
<i>Wenbo Huang, Jinghui Zhang, Guang Li, Lei Zhang, Shuoyuan Wang, Fang Dong, Jiahui Jin, Takahiro Ogawa, Miki Haseyama</i>	
CLIP-RestoreX: Restore Image Structure and Perception in Exposure Correction.....	3760
<i>Xiang Huang, Qing Zhang, Jian-Fang Hu, Wei-Shi Zheng</i>	
DAMPER: A Dual-Stage Medical Report Generation Framework with Coarse-Grained MeSH Alignment and Fine-Grained Hypergraph Matching	3769
<i>Xiaofei Huang, Wenting Chen, Jie Liu, Qisheng Lu, Xiaoling Luo, Linlin Shen</i>	
Towards a Multimodal Large Language Model with Pixel-Level Insight for Biomedicine	3779
<i>Xiaoshuang Huang, Lingdong Shen, Jia Liu, Fangxin Shang, Hongxiang Li, Haifeng Huang, Yehui Yang</i>	
PSReg: Prior-guided Sparse Mixture of Experts for Point Cloud Registration	3788
<i>Xiaoshui Huang, Zhou Huang, Yifan Zuo, Yongshun Gong, Chengdong Zhang, Deyang Liu, Yuming Fang</i>	
Medical MLLM Is Vulnerable: Cross-Modality Jailbreak and Mismatched Attacks on Medical Multimodal Large Language Models	3797
<i>Xijie Huang, Xinyuan Wang, Hantao Zhang, Yinghao Zhu, Jiawen Xi, Jingkun An, Hao Wang, Hao Liang, Chengwei Pan</i>	
L4DR: LiDAR-4DRadar Fusion for Weather-Robust 3D Object Detection	3806
<i>Xun Huang, Ziyu Xu, Hai Wu, Jinlong Wang, Qiming Xia, Yan Xia, Jonathan Li, Kyle Gao, Chenglu Wen, Cheng Wang</i>	
Zero-Shot Low-Light Image Enhancement via Latent Diffusion Models	3815
<i>Yan Huang, Xiaoshan Liao, Jinxiu Liang, Yuhui Quan, Boxin Shi, Yong Xu</i>	
Distilling Knowledge from Heterogeneous Architectures for Semantic Segmentation.....	3824
<i>Yanglin Huang, Kai Hu, Yuan Zhang, Zhineng Chen, Xieping Gao</i>	
SeFAR: Semi-supervised Fine-grained Action Recognition with Temporal Perturbation and Learning Stabilization	3833
<i>Yongle Huang, Haodong Chen, Zhenbang Xu, Zihan Jia, Haozhou Sun, Dian Shao</i>	
PoseMamba: Monocular 3D Human Pose Estimation with Bidirectional Global-Local Spatio-Temporal State Space Model.....	3842
<i>Yunlong Huang, Junshuo Liu, Ke Xian, Robert Caiming Qiu</i>	
Generative Medical Segmentation.....	3851
<i>Jiayu Huo, Xi Ouyang, Sébastien Ourselin, Rachel Sparks</i>	
EGSRAL:An Enhanced 3D Gaussian Splatting Based Renderer with Automated Labeling for Large-Scale Driving Scene	3860
<i>Yixiong Huo, Guangfeng Jiang, Hongyang Wei, Ji Liu, Song Zhang, Han Liu, Xingliang Huang, Mingjie Lu, Jinzhang Peng, Dong Li, Lu Tian, Emad Barsoum</i>	
High-Resolution Frame Interpolation with Patch-based Cascaded Diffusion	3868
<i>Junhwa Hur, Charles Herrmann, Saurabh Saxena, Janne Kontkanen, Wei-Sheng Lai, Yichang Shih, Michael Rubinstein, David J. Fleet, Deqing Sun</i>	

Zero-shot Depth Completion via Test-time Alignment with Affine-invariant Depth Prior	3877
<i>Lee Hyoseok, Kyeong Seon Kim, Kwon Byung-Ki, Tae-Hyun Oh</i>	
VG-TVP: Multimodal Procedural Planning via Visually Grounded Text-Video Prompting.....	3886
<i>Muhammet Furkan Ilaslan, Ali Köksal, Kevin Qinghong Lin, Burak Satar, Mike Zheng Shou, Qianli Xu</i>	
QORT-Former: Query-optimized Real-time Transformer for Understanding Two Hands Manipulating Objects	3895
<i>Elkhan Ismayilzada, MD Khalequzzaman Chowdhury Sayem, Yihalem Yimolal Tiruneh, Mubarrat Tajoar Chowdhury, Muhammadjon Boboev, Seungryul Baek</i>	
Every Component Counts: Rethinking the Measure of Success for Medical Semantic Segmentation in Multi-Instance Segmentation Tasks.....	3904
<i>Alexander Jaus, Constantin Marc Seibold, Simon Reiß, Zdravko Marinov, Keyi Li, Zeling Ye, Stefan Krieg, Jens Kleesiek, Rainer Stiefelwagen</i>	
Game4Loc: A UAV Geo-Localization Benchmark from Game Data	3913
<i>Yuxiang Ji, Boyong He, Zhuoyue Tan, Liaoni Wu</i>	
FastLGS: Speeding Up Language Embedded Gaussians with Feature Grid Mapping.....	3922
<i>Yuzhou Ji, He Zhu, Junshu Tang, Wuyi Liu, Zhizhong Zhang, Xin Tan, Yuan Xie</i>	
ContextHOI: Spatial Context Learning for Human-Object Interaction Detection	3931
<i>Mingda Jia, Liming Zhao, Ge Li, Yun Zheng</i>	
Orchestrating the Symphony of Prompt Distribution Learning for Human-Object Interaction Detection	3940
<i>Mingda Jia, Liming Zhao, Ge Li, Yun Zheng</i>	
Doubly Contrastive Learning for Source-Free Domain Adaptive Person Search.....	3949
<i>Yizhen Jia, Rong Quan, Yue Feng, Haiyan Chen, Jie Qin</i>	
DesignEdit: Unify Spatial-Aware Image Editing via Training-free Inpainting with a Multi-Layered Latent Diffusion Framework	3958
<i>Yueru Jia, Aosong Cheng, Yuhui Yuan, Chuke Wang, Ji Li, Huizhu Jia, Shanghang Zhang</i>	
FlexiTex: Enhancing Texture Generation via Visual Guidance	3967
<i>Dadong Jiang, Xianghui Yang, Zibo Zhao, Sheng Zhang, Jiaao Yu, Zeqiang Lai, Shaoxiong Yang, Chunchao Guo, Xiaobo Zhou, Zhihui Ke</i>	
Granularity-Adaptive Spatial Evidence Tokenization for Video Question Answering.....	3976
<i>Hao Jiang, Yang Jin, Zhicheng Sun, Kun Xu, Kun Xu, Liwei Chen, Yang Song, Kun Gai, Yadong Mu</i>	
ARNet: Self-Supervised FG-SBIR with Unified Sample Feature Alignment and Multi-Scale Token Recycling.....	3985
<i>Jianan Jiang, Hao Tang, Zhilin Jiang, Weiren Yu, Di Wu</i>	
RRT-MVS: Recurrent Regularization Transformer for Multi-View Stereo.....	3994
<i>Jianfei Jiang, Liyong Wang, Haochen Yu, Tianyu Hu, Jiansheng Chen, Huimin Ma</i>	

PART 2

SCCS: Deep Neural Spectral Clustering for Self-Supervised Subcellular Structure Segmentation	4003
<i>Jimao Jiang, Diya Sun, Tianbing Wang, Yuru Pei</i>	

PixelMan: Consistent Object Editing with Diffusion Models via Pixel Manipulation and Generation.....	4012
<i>Liyao Jiang, Negar Hassanpour, Mohammad Salameh, Mohammadreza Samadi, Jiao He, Fengyu Sun, Di Niu</i>	
Restabilizing Diffusion Models with Predictive Noise Fusion Strategy for Image Super-Resolution	4021
<i>Luoqian Jiang, Yong Guo, Bingna Xu, Haolin Pan, Jiezhong Cao, Wenbo Li, Jian Chen</i>	
LATTE: Improving Latex Recognition for Tables and Formulae with Iterative Refinement.....	4030
<i>Nan Jiang, Shanchao Liang, Chengxiao Wang, Jiannan Wang, Lin Tan</i>	
Move and Act: Enhanced Object Manipulation and Background Integrity for Image Editing.....	4039
<i>Pengfei Jiang, Mingbao Lin, Fei Chao</i>	
Energy-Guided Optimization for Personalized Image Editing with Pretrained Text-to-Image Diffusion Models.....	4048
<i>Rui Jiang, Xinghe Fu, Guangcong Zheng, Teng Li, Taiping Yao, Xi Li</i>	
Query Quantized Neural SLAM.....	4057
<i>Sijia Jiang, Jing Hua, Zhizhong Han</i>	
Sensing Surface Patches in Volume Rendering for Inferring Signed Distance Functions	4066
<i>Sijia Jiang, Tong Wu, Jing Hua, Zhizhong Han</i>	
What Kind of Visual Tokens Do We Need? Training-Free Visual Token Pruning for Multi-Modal Large Language Models from the Perspective of Graph	4075
<i>Yutao Jiang, Qiong Wu, Wenhao Lin, Wei Yu, Yiyi Zhou</i>	
Weighted Poisson-disk Resampling on Large-Scale Point Clouds.....	4084
<i>Xianhe Jiao, Chenlei Lv, Junli Zhao, Ran Yi, Yu-Hui Wen, Zhenkuan Pan, Zhongke Wu, Yong-Jin Liu</i>	
SpatioTemporal Learning for Human Pose Estimation in Sparsely-Labeled Videos	4093
<i>Yingying Jiao, Zhigang Wang, Sifan Wu, Shaojing Fan, Zhenguang Liu, Zhuoyue Xu, Zheqi Wu</i>	
Optimizing Human Pose Estimation Through Focused Human and Joint Regions	4102
<i>Yingying Jiao, Zhigang Wang, Zhenguang Liu, Shaojing Fan, Sifan Wu, Zheqi Wu, Zhuoyue Xu</i>	
Visual Prompting Upgrades Neural Network Sparsification: A Data-Model Perspective	4111
<i>Can Jin, Tianjin Huang, Yihua Zhang, Mykola Pechenizkiy, Sijia Liu, Shiwei Liu, Tianlong Chen</i>	
Exploring More from Multiple Gait Modalities for Human Identification.....	4120
<i>Dongyang Jin, Chao Fan, Weihua Chen, Shiqi Yu</i>	
LogicAD: Explainable Anomaly Detection via VLM-based Text Feature Extraction.....	4129
<i>Er Jin, Qihui Feng, Yongli Mou, Gerhard Lakemeyer, Stefan Decker, Oliver Simons, Johannes Stegmaier</i>	
Pedestrian Attribute Recognition: A New Benchmark Dataset and a Large Language Model Augmented Framework	4138
<i>Jiandong Jin, Xiao Wang, Qian Zhu, Haiyang Wang, Chenglong Li</i>	
A Method for Enhancing Generalization of Adam by Multiple Integrations.....	4147
<i>Long Jin, Han Nong, Liangming Chen, Zhenming Su</i>	

Constructing Fair Latent Space for Intersection of Fairness and Explainability	4156
<i>Hyungjun Joo, Hyeonggeun Han, Sehwan Kim, Sangwoo Hong, Jungwoo Lee</i>	
Bridging the Semantic Granularity Gap Between Text and Frame Representations for Partially Relevant Video Retrieval.....	4166
<i>WooJin Jun, WonJun Moon, Cheol-Ho Cho, MinSeok Jung, Jae-Pil Heo</i>	
Event-Enhanced Blurry Video Super-Resolution	4175
<i>Dachun Kai, Yueyi Zhang, Jin Wang, Zeyu Xiao, Zhiwei Xiong, Xiaoyan Sun</i>	
NeSyCoCo: A Neuro-Symbolic Concept Composer for Compositional Generalization.....	4184
<i>Danial Kamali, Elham J. Barezi, Parisa Kordjamshidi</i>	
Exploring Enhanced Contextual Information for Video-Level Object Tracking.....	4194
<i>Ben Kang, Xin Chen, Simiao Lai, Yang Liu, Yi Liu, Dong Wang</i>	
CodecNeRF: Toward Fast Encoding and Decoding, Compact, and High-quality Novel-view Synthesis.....	4203
<i>Gyeongjin Kang, Younggeun Lee, Seungjun Oh, Eunbyung Park</i>	
C2PD: Continuity-Constrained Pixelwise Deformation for Guided Depth Super-Resolution	4212
<i>Jiahui Kang, Qing Cai, Runqing Tan, Yimei Liu, Zhi Liu</i>	
DiffusionREC: Diffusion Model with Adaptive Condition for Referring Expression Comprehension.....	4221
<i>Jingcheng Ke, Waikeng Wong, Jia Wang, Mu Li, Lunke Fei, Jie Wen</i>	
Learning to Prompt with Text Only Supervision for Vision-Language Models	4230
<i>Muhammad Uzair Khattak, Muhammad Ferjad Naeem, Muzammal Naseer, Luc Van Gool, Federico Tombari</i>	
PLATYPUS: Progressive Local Surface Estimator for Arbitrary-Scale Point Cloud Upsampling	4239
<i>Donghyun Kim, Hyeonkyeong Kwon, Yumin Kim, Seong Jae Hwang</i>	
Generalized Zero-Shot Learning for Point Cloud Segmentation with Evidence-Based Dynamic Calibration.....	4248
<i>Hyeonseok Kim, Byeongkeun Kang, Yeejin Lee</i>	
APR-RD: Complemental Two Steps for Self-Supervised Real Image Denoising.....	4257
<i>Hyunjun Kim, Nam Ik Cho</i>	
Prediction-Feedback DETR for Temporal Action Detection.....	4266
<i>Jihwan Kim, Miso Lee, Cheol-Ho Cho, Jihyun Lee, Jae-Pil Heo</i>	
DEEPTalk: Dynamic Emotion Embedding for Probabilistic Speech-Driven 3D Face Animation.....	4275
<i>Jisoo Kim, Jungbin Cho, Joonho Park, Soonmin Hwang, Da Eun Kim, Geon Kim, Youngjae Yu</i>	
ProtoOcc: Accurate, Efficient 3D Occupancy Prediction Using Dual Branch Encoder-Prototype Query Decoder	4284
<i>Jungho Kim, Changwon Kang, Dongyoung Lee, Sehwan Choi, Jun Won Choi</i>	
HiCM ² : Hierarchical Compact Memory Modeling for Dense Video Captioning.....	4293
<i>Minkuk Kim, Hyeon Bae Kim, Jinyoung Moon, Jinwoo Choi, Seong Tae Kim</i>	
MoDiTalker: Motion-Disentangled Diffusion Model for High-Fidelity Talking Head Generation.....	4302
<i>Seyeon Kim, Siyoon Jin, Jihye Park, Kihong Kim, Jiyoung Kim, Jisu Nam, Seungryong Kim</i>	

TSDf-Based Efficient Motion-Compensated Temporal Interpolation for 3D Dynamic Sequences	4311
<i>Soowoong Kim, Minseong Kwon, Junho Choi, Gun Bang, Seungjoon Yang</i>	
ViPCap: Retrieval Text-Based Visual Prompts for Lightweight Image Captioning	4320
<i>Taewhan Kim, Soeun Lee, Si-Woo Kim, Dong-Jin Kim</i>	
Multi-Modal Grounded Planning and Efficient Replanning for Learning Embodied Agents with a Few Examples	4329
<i>Taewoong Kim, Byeonghwi Kim, Jonghyun Choi</i>	
DiffuseHigh: Training-Free Progressive High-Resolution Image Synthesis Through Structure Guidance.....	4338
<i>Younghyun Kim, Geunmin Hwang, Junyu Zhang, Eunbyung Park</i>	
SatCLIP: Global, General-Purpose Location Embeddings with Satellite Imagery	4347
<i>Konstantin Klemmer, Esther Rolf, Caleb Robinson, Lester Mackey, Marc Rußwurm</i>	
Sequence Matters: Harnessing Video Models in 3D Super-Resolution.....	4356
<i>Hyun-kyu Ko, Dongheok Park, Youngin Park, Byeonghyeon Lee, Juhee Han, Eunbyung Park</i>	
UniDet3D: Multi-dataset Indoor 3D Object Detection.....	4365
<i>Maksim Kolodiaznyy, Anna Vorontsova, Matvey Skripkin, Danila Rukhovich, Anton Konushin</i>	
Efficient Gaussian Splatting for Monocular Dynamic Scene Rendering via Sparse Time-Variant Attribute Modeling	4374
<i>Hanyang Kong, Xingyi Yang, Xinchao Wang</i>	
Do Not DeepFake Me: Privacy-Preserving Neural 3D Head Reconstruction Without Sensitive Images	4383
<i>Jiayi Kong, Xurui Song, Shuo Huai, Baixin Xu, Jun Luo, Ying He</i>	
Real-Time Neural Denoising with Render-Aware Knowledge Distillation.....	4392
<i>Mengxun Kong, Jie Guo, Chen Wang, Ye Yuan, Yanwen Guo</i>	
MHBench: Demystifying Motion Hallucination in VideoLLMs.....	4401
<i>Ming Kong, Xianzhou Zeng, Luyuan Chen, Yadong Li, Bo Yan, Qiang Zhu</i>	
COLUMBUS: Evaluating COgnitive Lateral Understanding Through Multiple-Choice reBUSes	4410
<i>Koen Kraaijveld, Yifan Jiang, Kaixin Ma, Filip Ilievski</i>	
Stable Mean Teacher for Semi-supervised Video Action Detection	4419
<i>Akash Kumar, Sirshapan Mitra, Yogesh Singh Rawat</i>	
A Unified Degradation-Robust Approach to SSL and UDA for 3D Medical Images.....	4428
<i>Suruchi Kumari, Pravendra Singh</i>	
SAFIRE: Segment Any Forged Image Region.....	4437
<i>Myung-Joon Kwon, Wonjun Lee, Seung-Hun Nam, Minji Son, Changick Kim</i>	
Mind the Uncertainty in Human Disagreement: Evaluating Discrepancies Between Model Predictions and Human Responses in VQA	4446
<i>Jian Lan, Diego Frassinelli, Barbara Plank</i>	
Exploiting Diffusion Prior for Real-World Image Dehazing with Unpaired Training.....	4455
<i>Yunwei Lan, Zhigao Cui, Chang Liu, Jialun Peng, Nian Wang, Xin Luo, Dong Liu</i>	

Color Transfer with Modulated Flows.....	4464
<i>Maria Larchenko, Alexander Lobashev, Dmitry Guskov, Vladimir Vladimirovich Palyulin</i>	
Progressive Multi-granular Alignments for Grounded Reasoning in Large Vision-Language Models	4473
<i>Quang-Hung Le, Long Hoang Dang, Ngan Hoang Le, Truyen Tran, Thao Minh Le</i>	
Multispectral Pedestrian Detection with Sparsely Annotated Label.....	4482
<i>Chan Lee, Seungho Shin, Gyeong-Moon Park, Jung Uk Kim</i>	
Rethinking Open-Vocabulary Segmentation of Radiance Fields in 3D Space	4491
<i>Hyunjee Lee, Youngsik Yun, Jeongmin Bae, Seoha Kim, Youngjung Uh</i>	
VidChain: Chain-of-Tasks with Metric-based Direct Preference Optimization for Dense Video Captioning	4499
<i>Ji Soo Lee, Jongha Kim, Jeehye Na, Jinyoung Park, Hyunwoo J. Kim</i>	
NBA3D: Neighbor-Based Confidence Adjustment for 3D Rare Object Detection Using LiDAR.....	4508
<i>Jooyoung Lee, Jaeyoon Lee, Jongwon Choi</i>	
Navigating Label Ambiguity for Facial Expression Recognition in the Wild	4517
<i>JunGyu Lee, Yeji Choi, Haksub Kim, Ig-Jae Kim, Gi Pyo Nam</i>	
Video Diffusion Models Are Strong Video Inpainter	4526
<i>Minhyeok Lee, Suhwan Cho, Chajin Shin, Jungho Lee, Sunghun Yang, Sangyoun Lee</i>	

Author Index