

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 3: AAAI Technical Tracks

- Computer Vision II

Philadelphia, Pennsylvania, USA
25 February - 4 March 2025

Part 1 of 2

ISBN: 979-8-3313-2513-8

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 II

SVGBuilder: Component-Based Colored SVG Generation with Text-Guided Autoregressive Transformers.....	2358
<i>Zehao Chen, Rong Pan</i>	
EvHDR-GS: Event-guided HDR Video Reconstruction with 3D Gaussian Splatting.....	2367
<i>Zehao Chen, Zhan Lu, De Ma, Huajin Tang, Xudong Jiang, Qian Zheng, Gang Pan</i>	
EvHDR-NeRF: Building High Dynamic Range Radiance Fields with Single Exposure Images and Events	2376
<i>Zehao Chen, Zhanfeng Liao, De Ma, Huajin Tang, Qian Zheng, Gang Pan</i>	
VFM-Adapter: Adapting Visual Foundation Models for Dense Prediction with Dynamic Hybrid Operation Mapping.....	2385
<i>Zheng Chen, Yu Zeng, Zehui Chen, Hongzhi Gao, Lin Chen, Jiaming Liu, Feng Zhao</i>	
VersaGen: Unleashing Versatile Visual Control for Text-to-Image Synthesis	2394
<i>Zhipeng Chen, Lan Yang, Yonggang Qi, Honggang Zhang, Kaiyue Pang, Ke Li, Yi-Zhe Song</i>	
EchoMimic: Lifelike Audio-Driven Portrait Animations through Editable Landmark Conditions	2403
<i>Zhiyuan Chen, Jiajiong Cao, Zhiquan Chen, Yuming Li, Chenguang Ma</i>	
Spatiotemporal Blind-Spot Network with Calibrated Flow Alignment for Self-Supervised Video Denoising	2411
<i>Zikang Chen, Tao Jiang, Xiaowan Hu, Wang Zhang, Huaqiu Li, Haoqian Wang</i>	
Filter or Compensate: Towards Invariant Representation from Distribution Shift for Anomaly Detection	2420
<i>Zining Chen, Xingshuang Luo, Weiqiu Wang, Zhicheng Zhao, Fei Su, Aidong Men</i>	
Gradient Alignment Improves Test-Time Adaptation for Medical Image Segmentation	2429
<i>Ziyang Chen, Yiwen Ye, Yongsheng Pan, Yong Xia</i>	
ResAdapter: Domain Consistent Resolution Adapter for Diffusion Models	2438
<i>Jiaxiang Cheng, Pan Xie, Xin Xia, Jiashi Li, Jie Wu, Yuxi Ren, Huixia Li, Xuefeng Xiao, Shilei Wen, Lean Fu</i>	
3DPGS: 3D Probabilistic Graph Search for Archaeological Piece Grouping.....	2447
<i>Junfeng Cheng, Yingkai Yang, Tania Stathaki</i>	
Effective Diffusion Transformer Architecture for Image Super-Resolution.....	2455
<i>Kun Cheng, Lei Yu, Zhijun Tu, Xiao He, Liyu Chen, Yong Guo, Mingrui Zhu, Nannan Wang, Xinbo Gao, Jie Hu</i>	
DIDiffGes: Decoupled Semi-Implicit Diffusion Models for Real-time Gesture Generation from Speech	2464
<i>Yongkang Cheng, Shaoli Huang, Xuelin Chen, Jifeng Ning, Mingming Gong</i>	

Graphic Design with Large Multimodal Model	2473
<i>Yutao Cheng, Zhao Zhang, Maoke Yang, Hui Nie, Chunyuan Li, Xinglong Wu, Jie Shao</i>	
Aligning Instance Brownian Bridge with Texts for Open-Vocabulary Video Instance Segmentation.....	2482
<i>Zesen Cheng, Kehan Li, Li Hao, Peng Jin, Xiawu Zheng, Chang Liu, Jie Chen</i>	
Bridge 2D-3D: Uncertainty-aware Hierarchical Registration Network with Domain Alignment	2491
<i>Zhixin Cheng, Jiacheng Deng, Xinjun Li, Baoqun Yin, Tianzhu Zhang</i>	
Ambiguity-Restrained Text-Video Representation Learning for Partially Relevant Video Retrieval	2500
<i>Cheol-Ho Cho, WonJun Moon, WooJin Jun, MinSeok Jung, Jae-Pil Heo</i>	
Zero-Shot Scene Change Detection.....	2509
<i>Kyusik Cho, Dong Yeop Kim, Euntai Kim</i>	
Enhancing Robustness in Incremental Learning with Adversarial Training.....	2518
<i>Seungju Cho, Hongsin Lee, Changick Kim</i>	
Elevating Flow-Guided Video Inpainting with Reference Generation	2527
<i>Suhwan Cho, Seoung Wug Oh, Sangyoun Lee, Joon-Young Lee</i>	
Distribution-Level Feature Distancing for Machine Unlearning: Towards a Better Trade-off Between Model Utility and Forgetting	2536
<i>Dasol Choi, Dongbin Na</i>	
SIDL: A Real-World Dataset for Restoring Smartphone Images with Dirty Lenses	2545
<i>Sooyoung Choi, Sungyong Park, Heewon Kim</i>	
Intrinsic Image Decomposition for Robust Self-supervised Monocular Depth Estimation on Reflective Surfaces	2555
<i>Wonhyeok Choi, Kyumin Hwang, Minwoo Choi, Kiljoon Han, Wonjoon Choi, Mingyu Shin, Sunghoon Im</i>	
DynASyn: Multi-Subject Personalization Enabling Dynamic Action Synthesis.....	2564
<i>Yongjin Choi, Chanhun Park, Seung Jun Baek</i>	
Digging into Intrinsic Contextual Information for High-fidelity 3D Point Cloud Completion	2573
<i>Jisheng Chu, Wenrui Li, Xingtao Wang, Kanglin Ning, Yidan Lu, Xiaopeng Fan</i>	
What to Preserve and What to Transfer: Faithful, Identity-Preserving Diffusion-based Hairstyle Transfer	2582
<i>Chaeyeon Chung, Sunghyun Park, Jeongho Kim, Jaegul Choo</i>	
MASS: Overcoming Language Bias in Image-Text Matching.....	2591
<i>Jiwan Chung, Seungwon Lim, Sangkyu Lee, Youngjae Yu</i>	
AttackBench: Evaluating Gradient-based Attacks for Adversarial Examples	2600
<i>Antonio Emanuele Cinà, Jérôme Rony, Maura Pintor, Luca Demetrio, Ambra Demontis, Battista Biggio, Ismail Ben Ayed, Fabio Roli</i>	
LOMA: Language-assisted Semantic Occupancy Network via Triplane Mamba	2609
<i>Yubo Cui, Zhiheng Li, Jiaqiang Wang, Zheng Fang</i>	
Multi-task Visual Grounding with Coarse-to-Fine Consistency Constraints.....	2618
<i>Ming Dai, Jian Li, Jiedong Zhuang, Xian Zhang, Wankou Yang</i>	

DCSF-KD: Dynamic Channel-wise Spatial Feature Knowledge Distillation for Object Detection.....	2627
<i>Tao Dai, Yang Lin, Hang Guo, Jinbao Wang, Zexuan Zhu</i>	
GCD-Sampling: A General Cross-scale Decoupled Sampling for Point Cloud	2636
<i>Tao Dai, Yanzi Wang, Jianyu Xiong, Yaohua Zha, Shu-Tao Xia, Zexuan Zhu</i>	
Harmonious Music-driven Group Choreography with Trajectory-Controllable Diffusion	2645
<i>Yuqin Dai, Wanlu Zhu, Ronghui Li, Zeping Ren, Xiangzheng Zhou, Jixuan Ying, Jun Li, Jian Yang</i>	
Self-Corrected Flow Distillation for Consistent One-Step and Few-Step Image Generation.....	2654
<i>Quan Dao, Hao Phung, Trung Tuan Dao, Dimitris N. Metaxas, Anh Tran</i>	
PIXELS: Progressive Image Exemplar-based Editing with Latent Surgery	2663
<i>Shristi Das Biswas, Matthew Shreve, Xuelu Li, Prateek Singhal, Kaushik Roy</i>	
Single Exposure Quantitative Phase Imaging with a Conventional Microscope Using Diffusion Models.....	2672
<i>Gabriel della Maggiora, Luis Alberto Croquevielle, Harry Horsley, Thomas Heinis, Artur Yakimovich</i>	
Deep Non-Rigid Structure-from-Motion Revisited: Canonicalization and Sequence Modeling.....	2681
<i>Hui Deng, Jiawei Shi, Zhen Qin, Yiran Zhong, Yuchao Dai</i>	
DiffCorr: Conditional Diffusion Model with Reliable Pseudo-Label Guidance for Unsupervised Point Cloud Shape Correspondence	2690
<i>Jiacheng Deng, Jiahao Lu, Zhixin Cheng, Wenfei Yang</i>	
Adaptive Siamese Masked Autoencoder with Global Optimization for Unsupervised Point Cloud Shape Correspondence	2699
<i>Jiacheng Deng, Jiahao Lu</i>	
OTIAS: OcTree Implicit Adaptive Sampling for Multispectral and Hyperspectral Image Fusion.....	2708
<i>Shangqi Deng, Jun Ma, Liang-Jian Deng, Ping Wei</i>	
Boundary-Aware Temporal Dynamic Pseudo-Supervision Pairs Generation for Zero-Shot Natural Language Video Localization.....	2717
<i>Xiongwen Deng, Haoyu Tang, Han Jiang, Qinghai Zheng, Jihua Zhu</i>	
Occlusion-Insensitive Talking Head Video Generation via Facelet Compensation.....	2726
<i>Yuhui Deng, Yuqin Lu, Yangyang Xu, Yongwei Nie, Shengfeng He</i>	
SSLFusion: Scale and Space Aligned Latent Fusion Model for Multimodal 3D Object Detection	2735
<i>Bonan Ding, Jin Xie, Jing Nie, Jiale Cao</i>	
Dis ² Booth: Learning Image Distribution with Disentangled Features for Text-to-Image Diffusion Models.....	2744
<i>Guanqi Ding, Chengyu Yang, Shuhui Wang, Xincheng Li, Jinzhe Zhang, Xin Jin, Qingming Huang</i>	
Muses: 3D-Controllable Image Generation via Multi-Modal Agent Collaboration	2753
<i>Yanbo Ding, Shaobin Zhuang, Kunchang Li, Zhengrong Yue, Yu Qiao, Yali Wang</i>	
AS-Det: Active Sampling for Adaptive 3D Object Detection in Point Clouds.....	2762
<i>Ziheng Ding, Xiaze Zhang, Qi Jing, Ying Cheng, Rui Feng</i>	

GarFast: Realistic and Fast Garment Transfer with a Simplified Parser-Free Approach.....	2771
<i>Chenghu Du, Junyin Wang, Yi Rong, Feng Yu, Shengwu Xiong</i>	
Latent Diffusion-Enhanced Virtual Try-On via Optimized Pseudo-Label Generation.....	2780
<i>Chenghu Du, Junyin Wang, Feng Yu, Shengwu Xiong</i>	
HybridReg: Robust 3D Point Cloud Registration with Hybrid Motions	2789
<i>Keyu Du, Hao Xu, Haipeng Li, Hong Qu, Chi-Wing Fu, Shuaicheng Liu</i>	
Out of Length Text Recognition with Sub-String Matching.....	2798
<i>Yongkun Du, Zhineng Chen, Caiyan Jia, Xieping Gao, Yu-Gang Jiang</i>	
InstructOCR: Instruction Boosting Scene Text Spotting	2807
<i>Chen Duan, Qianyi Jiang, Pei Fu, Jiamin Chen, Shengxi Li, Zining Wang, Shan Guo, Junfeng Luo</i>	
A Diffusion-Based Framework for Occluded Object Movement	2816
<i>Zheng-Peng Duan, Jiawei Zhang, Siyu Liu, Zheng Lin, Chun-Le Guo, Dongqing Zou, Jimmy Ren, Chongyi Li</i>	
DiffRetouch: Using Diffusion to Retouch on the Shoulder of Experts	2825
<i>Zheng-Peng Duan, Jiawei Zhang, Zheng Lin, Xin Jin, XunDong Wang, Dongqing Zou, Chun-Le Guo, Chongyi Li</i>	
IniRetinex: Rethinking Retinex-type Low-Light Image Enhancer via Initialization Perspective	2834
<i>Guodong Fan, Zishu Yao, Guang-Yong Chen, Jian-Nan Su, Min Gan</i>	
Vision-guided Text Mining for Unsupervised Cross-modal Hashing with Community Similarity Quantization	2843
<i>Haozhi Fan, Yuan Cao</i>	
Depth-Centric Dehazing and Depth-Estimation from Real-World Hazy Driving Video.....	2852
<i>Junkai Fan, Kun Wang, Zhiqiang Yan, Xiang Chen, Shangbing Gao, Jun Li, Jian Yang</i>	
EventPillars: Pillar-based Efficient Representations for Event Data.....	2861
<i>Rui Fan, Weidong Hao, Juntao Guan, Lai Rui, Lin Gu, Tong Wu, Fanhong Zeng, Zhangming Zhu</i>	
Combating Semantic Contamination in Learning with Label Noise	2870
<i>Wenxiao Fan, Kan Li</i>	
EMHI: A Multimodal Egocentric Human Motion Dataset with HMD and Body-Worn IMUs	2879
<i>Zhen Fan, Peng Dai, Zhuo Su, Xu Gao, Zheng Lv, Jiarui Zhang, Tianyuan Du, Guidong Wang, Yang Zhang</i>	
CoSDA: Enhancing the Robustness of Inversion-based Generative Image Watermarking Framework.....	2888
<i>Han Fang, Kejiang Chen, Zijin Yang, Bosen Cui, Weiming Zhang, Ee-Chien Chang</i>	
SSUN-Net: Spatial-Spectral Prior-Aware Unfolding Network for Pan-Sharpening.....	2897
<i>Shijie Fang, Hongping Gan</i>	
Guided Real Image Dehazing Using YCbCr Color Space	2906
<i>Wenxuan Fang, Junkai Fan, Yu Zheng, Jiangwei Weng, Ying Tai, Jun Li</i>	

PART 2

Multi-Pair Temporal Sentence Grounding via Multi-Thread Knowledge Transfer Network	2915
<i>Xiang Fang, Wanlong Fang, Changshuo Wang, Daizong Liu, Keke Tang, Jianfeng Dong, Pan Zhou, Beibei Li</i>	
AE-NeRF: Augmenting Event-Based Neural Radiance Fields for Non-ideal Conditions and Larger Scenes.....	2924
<i>Chaoran Feng, Wangbo Yu, Xinhua Cheng, Zhenyu Tang, Junwu Zhang, Li Yuan, Yonghong Tian</i>	
PROSAC: Provably Safe Certification for Machine Learning Models under Adversarial Attacks	2933
<i>Chen Feng, Ziquan Liu, Zhuo Zhi, Ilija Bogunovic, Carsten Gerner-Beuerle, Miguel Rodrigues</i>	
VQA4CIR: Boosting Composed Image Retrieval with Visual Question Answering	2942
<i>Chun-Mei Feng, Yang Bai, Tao Luo, Zhen Li, Salman Khan, Wangmeng Zuo, Rick Siow Mong Goh, Yong Liu</i>	
PoseLLaVA: Pose Centric Multimodal LLM for Fine-Grained 3D Pose Manipulation.....	2951
<i>Dong Feng, Ping Guo, Encheng Peng, Mingmin Zhu, Wenhao Yu, Peng Wang</i>	
Residual Diffusion Deblurring Model for Single Image Defocus Deblurring.....	2960
<i>Haoxuan Feng, Haohui Zhou, Tian Ye, Sixiang Chen, Lei Zhu</i>	
DiT4Edit: Diffusion Transformer for Image Editing.....	2969
<i>Kunyu Feng, Yue Ma, Bingyuan Wang, Chenyang Qi, Haozhe Chen, Qifeng Chen, Zeyu Wang</i>	
Semantic Ambiguity Modeling and Propagation for Fine-Grained Visual Cross View Geo-Localization	2978
<i>Mingtao Feng, Fenghao Tian, Jianqiao Luo, Zijie Wu, Weisheng Dong, Yaonan Wang, Ajmal Saeed Mian</i>	
Weakly Supervised Gland Segmentation with Class Semantic Consistency and Purified Labels Filtration	2987
<i>Siyang Feng, Huadeng Wang, Chu Han, Zhenbing Liu, Hualong Zhang, Rushi Lan, Xipeng Pan</i>	
HDLayout: Hierarchical and Directional Layout Planning for Arbitrary Shaped Visual Text Generation	2996
<i>Tonghui Feng, Chunsheng Yan, Qianru Wang, Jiangtao Cui, Xiaotian Qiao</i>	
ViPOcc: Leveraging Visual Priors from Vision Foundation Models for Single-View 3D Occupancy Prediction	3004
<i>Yi Feng, Yu Han, Xijing Zhang, Tanghui Li, Yanting Zhang, Rui Fan</i>	
Simplifying Control Mechanism in Text-to-Image Diffusion Models.....	3013
<i>Zhida Feng, Li Chen, Yuenan Sun, Jiaxiang Liu, Shikun Feng</i>	
BGHR: Bridging the Gap Between HBox-Supervised and RBox-Supervised Oriented Object Detection via Adaptive Fine-Grained Sample Mining.....	3022
<i>Chenlin Fu, Yingying Zhu</i>	
Foundation Model Driven Appearance Extraction for Robust Multiple Object Tracking	3031
<i>Teng Fu, Haiyang Yu, Ke Niu, Bin Li, Xiangyang Xue</i>	

Exploring Unbiased Deepfake Detection via Token-Level Shuffling and Mixing	3040
<i>Xinghe Fu, Zhiyuan Yan, Taiping Yao, Shen Chen, Xi Li</i>	
MFL-Owner: Ownership Protection for Multi-modal Federated Learning via Orthogonal Transform Watermark	3049
<i>Keke Gai, Dongjue Wang, Jing Yu, Mohan Wang, Liehuang Zhu, Qi Wu</i>	
DFDNet: Disentangling and Filtering Dynamics for Enhanced Video Prediction	3059
<i>Lianqiang Gan, Junyu Lai, Jingze Ju, Lianli Gao, Yi Bin</i>	
PNVC: Towards Practical INR-based Video Compression	3068
<i>Ge Gao, Ho Man Kwan, Fan Zhang, David Bull</i>	
AIM: Let Any Multimodal Large Language Models Embrace Efficient In-Context Learning	3077
<i>Jun Gao, Qian Qiao, Tianxiang Wu, Zili Wang, Ziqiang Cao, Wenjie Li</i>	
TC-LLaVA: Rethinking the Transfer of LLaVA from Image to Video Understanding with Temporal Considerations	3086
<i>Mingze Gao, Jingyu Liu, Mingda Li, Jiangtao Xie, Qingbin Liu, Kevin Zhao, Xi Chen, Hui Xiong</i>	
Learning 2D Invariant Affordance Knowledge for 3D Affordance Grounding.....	3095
<i>Xianqiang Gao, Pingrui Zhang, Delin Qu, Dong Wang, Zhigang Wang, Yan Ding, Bin Zhao</i>	
EventMamba: Enhancing Spatio-Temporal Locality with State Space Models for Event-Based Video Reconstruction	3104
<i>Chengjie Ge, Xueyang Fu, Peng He, Kunyu Wang, Chengzhi Cao, Zheng-Jun Zha</i>	
Implicit Location-Caption Alignment via Complementary Masking for Weakly-Supervised Dense Video Captioning.....	3113
<i>Shiping Ge, Qiang Chen, Zhiwei Jiang, Yafeng Yin, Liu Qin, Ziyao Chen, Qing Gu</i>	
ParseCaps: An Interpretable Parsing Capsule Network for Medical Image Diagnosis	3122
<i>Xinyu Geng, Jiaming Wang, Xiaolin Huang, Fanglin Chen, Jun Xu</i>	
Auto-Regressive Diffusion for Generating 3D Human-Object Interactions.....	3131
<i>Zichen Geng, Zeeshan Hayder, Wei Liu, Ajmal Saeed Mian</i>	
Domain Generalized Medical Landmark Detection via Robust Boundary-Aware Pre-Training.....	3140
<i>Haifan Gong, Yu Lu, Xiang Wan, Haofeng Li</i>	
Rethinking Masked Data Reconstruction Pretraining for Strong 3D Action Representation Learning.....	3149
<i>Tao Gong, Qi Chu, Bin Liu, Nenghai Yu</i>	
Queryable Prototype Multiple Instance Learning with Vision-Language Models for Incremental Whole Slide Image Classification	3158
<i>Jiaxiang Gou, Luping Ji, Pei Liu, Mao Ye</i>	
Efficient Connectivity-Preserving Instance Segmentation with Supervoxel-Based Loss Function	3167
<i>Anna Grim, Jayaram Chandrashekar, Uygur Sümbül</i>	
MaintainAvatar: A Maintainable Avatar Based on Neural Radiance Fields by Continual Learning.....	3176
<i>Shengbo Gu, Yu-Kun Qiu, Yu-Ming Tang, Ancong Wu, Wei-Shi Zheng</i>	
HGSFusion: Radar-Camera Fusion with Hybrid Generation and Synchronization for 3D Object Detection	3185
<i>Zijian Gu, Jianwei Ma, Yan Huang, Honghao Wei, Zhanye Chen, Hui Zhang, Wei Hong</i>	

OT-StainNet: Optimal Transport Driven Semantic Matching for Weakly Paired H&E-to-IHC Stain Transfer	3194
<i>Xianchao Guan, Yifeng Wang, Ye Zhang, Zheng Zhang, Yongbing Zhang</i>	
DepthFM: Fast Generative Monocular Depth Estimation with Flow Matching.....	3203
<i>Ming Gui, Johannes Schusterbauer, Ulrich Prestel, Pingchuan Ma, Dmytro Kotovenko, Olga Grebenkova, Stefan Andreas Baumann, Vincent Tao Hu, Björn Ommer</i>	
You Should Learn to Stop Denoising on Point Clouds in Advance.....	3212
<i>Chuchen Guo, Weijie Zhou, Zheng Liu, Ying He</i>	
Surgical Workflow Recognition and Blocking Effectiveness Detection in Laparoscopic Liver Resection with Pringle Maneuver.....	3220
<i>Diandian Guo, Weixin Si, Zhixi Li, Jialun Pei, Pheng-Ann Heng</i>	
Cross-Spectral Gaussian Splatting with Spatial Occupancy Consistency	3229
<i>Haipeng Guo, Huanyu Liu, Jiazheng Wen, Junbao Li</i>	
Enhancing Low-Rank Adaptation with Recoverability-Based Reinforcement Pruning for Object Counting	3238
<i>Haojie Guo, Junyu Gao, Yuan Yuan</i>	
Towards a Comprehensive, Efficient and Promptable Anatomic Structure Segmentation Model Using 3D Whole-Body CT Scans.....	3247
<i>Heng Guo, Jianfeng Zhang, Jiaxing Huang, Tony C. W. Mok, Dazhou Guo, Ke Yan, Le Lu, Dakai Jin, Minfeng Xu</i>	
MetaNeRV: Meta Neural Representations for Videos with Spatial-Temporal Guidance	3257
<i>Jialong Guo, Ke Liu, Jiangchao Yao, Zhihua Wang, Jiajun Bu, Haishuai Wang</i>	
PromptDet: A Lightweight 3D Object Detection Framework with LiDAR Prompts	3266
<i>Kun Guo, Qiang Ling</i>	
OpenVIS: Open-vocabulary Video Instance Segmentation.....	3275
<i>Pinxue Guo, Hao Huang, Peiyang He, Xuefeng Liu, Tianjun Xiao, Wenqiang Zhang</i>	
Controllable 3D Dance Generation Using Diffusion-Based Transformer U-Net	3284
<i>Puyuan Guo, Tuo Hao, Wenxin Fu, Yingming Gao, Ya Li</i>	
SpikeGS: Reconstruct 3D Scene Captured by a Fast-Moving Bio-Inspired Camera	3293
<i>Yijia Guo, Liwen Hu, Yuanxi Bai, Jiawei Yao, Lei Ma, Tiejun Huang</i>	
VTG-LLM: Integrating Timestamp Knowledge into Video LLMs for Enhanced Video Temporal Grounding.....	3302
<i>Yongxin Guo, Jingyu Liu, Mingda Li, Dingxin Cheng, Xiaoying Tang, Dianbo Sui, Qingbin Liu, Xi Chen, Kevin Zhao</i>	
LLaVA Needs More Knowledge: Retrieval Augmented Natural Language Generation with Knowledge Graph for Explaining Thoracic Pathologies	3311
<i>Ameer Hamza, Abdullah , Yong Hyun Ahn, Sungyoung Lee, Seong Tae Kim</i>	
DuMo: Dual Encoder Modulation Network for Precise Concept Erasure.....	3320
<i>Feng Han, Kai Chen, Chao Gong, Zhipeng Wei, Jingjing Chen, Yu-Gang Jiang</i>	
GGs: Generalizable Gaussian Splatting for Lane Switching in Autonomous Driving.....	3329
<i>Huasong Han, Kaixuan Zhou, Xiaoxiao Long, Yusen Wang, Chunxia Xiao</i>	

ProPose: Probabilistic 3D Human Pose Estimation with Instance-Level Distribution and Normalizing Flow.....	3338
<i>Jumin Han, Jun-Hee Kim, Seong-Whan Lee</i>	
DME-Driver: Integrating Human Decision Logic and 3D Scene Perception in Autonomous Driving	3347
<i>Wencheng Han, Dongqian Guo, Cheng-Zhong Xu, Jianbing Shen</i>	
Boosting Segment Anything Model Towards Open-Vocabulary Learning.....	3356
<i>Xumeng Han, Longhui Wei, Xuehui Yu, Zhiyang Dou, Xin He, Kuiran Wang, Yingfei Sun, Zhenjun Han, Qi Tian</i>	
CoDTS: Enhancing Sparsely Supervised Collaborative Perception with a Dual Teacher-Student Framework.....	3366
<i>Yushan Han, Hui Zhang, Honglei Zhang, Jing Wang, Yidong Li</i>	
AsyncDSB: Schedule-Asynchronous Diffusion Schrödinger Bridge for Image Inpainting.....	3374
<i>Zihao Han, Baoquan Zhang, Lisai Zhang, Shanshan Feng, Kenghong Lin, Guotao Liang, Yunming Ye, Joeq , Kolaye</i>	
ID-Sculpt: ID-aware 3D Head Generation from Single In-the-wild Portrait Image.....	3383
<i>Jinkun Hao, Junshu Tang, Jiangning Zhang, Ran Yi, Yijia Hong, Moran Li, Weijian Cao, Yating Wang, Chengjie Wang, Lizhuang Ma</i>	
Multi-Frame Deformable Look-Up Table for Compressed Video Quality Enhancement	3392
<i>Gang He, Guancheng Quan, Chang Wu, Shihao Wang, Dajiang Zhou, Yunsong Li</i>	
V2C-CBM: Building Concept Bottlenecks with Vision-to-Concept Tokenizer	3401
<i>Hangzhou He, Lei Zhu, Xinliang Zhang, Shuang Zeng, Qian Chen, Yanye Lu</i>	
PointRWKV: Efficient RWKV-Like Model for Hierarchical Point Cloud Learning.....	3410
<i>Qingdong He, Jiangning Zhang, Jinlong Peng, Haoyang He, Xiangtai Li, Yabiao Wang, Chengjie Wang</i>	
Efficient Online Training for Zero-Shot Time-Lapse Microscopy Denoising and Super-Resolution	3419
<i>Ruian He, Ri Cheng, Xinkai Lyu, Weimin Tan, Bo Yan</i>	
DiffCalib: Reformulating Monocular Camera Calibration as Diffusion-Based Dense Incident Map Generation	3428
<i>Xiankang He, Guangkai Xu, Bo Zhang, Hao Chen, Ying Cui, Dongyan Guo</i>	
MagicMan: Generative Novel View Synthesis of Humans with 3D-Aware Diffusion and Iterative Refinement	3437
<i>Xu He, Zhiyong Wu, Xiaoyu Li, Di Kang, Chaopeng Zhang, Jiangnan Ye, Liyang Chen, Xiangjun Gao, Han Zhang, Haolin Zhuang</i>	
Long-Tailed Out-of-Distribution Detection: Prioritizing Attention to Tail.....	3446
<i>Yina He, Lei Peng, Yongcun Zhang, Juanjuan Weng, Shaozi Li, Zhiming Luo</i>	
Achieving Speed-Accuracy Balance in Vision-based 3D Occupancy Prediction via Geometric-Semantic Disentanglement	3455
<i>Yulin He, Wei Chen, Siqi Wang, Tianci Xun, Yusong Tan</i>	
Disentangle Nighttime Lens Flares: Self-supervised Generation-based Lens Flare Removal	3464
<i>Yuwen He, Wei Wang, Wanyu Wu, Kui Jiang</i>	

Author Index