

40th AAAI Conference on Artificial Intelligence (AAAI-26), 38th Conference on Innovative Applications of Artificial Intelligence (IAAI-26), and 16th Symposium on Educational Advances in Artificial Intelligence (EAAI-26)

Volume 16: AAAI Technical Tracks

- Computer Vision XIII

Singapore
20-27 January 2026

ISBN: 979-8-3313-3390-4

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© (2026) by Association for the Advancement of Artificial Intelligence
All rights reserved.

Printed with permission by Curran Associates, Inc. (2026)

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

AAAI TECHNICAL TRACK ON COMPUTER VISION XIII

CLIPDet3D: Vision-Language Collaborative Distillation for 3D Object Detection	13154
<i>Jiaqi Zhao, Huanfeng Hu, Yong Zhou, Wen-Liang Du, Kunyang Sun, Rui Yao, Qigong Sun</i>	
Facial Dynamics in Video: Instruction Tuning for Improved Facial Expression Perception and Contextual Awareness.....	13163
<i>Jiaxing Zhao, Boyuan Sun, Xiang Chen, Xihan Wei</i>	
Learning Procedural-Aware Video Representations Through State-Grounded Hierarchy Unfolding	13172
<i>Jinghan Zhao, Yifei Huang, Feng Lu</i>	
Beyond Predictive Resampling: Learning Input-Agnostic Downsampling for Efficient Aligned Vision Recognition	13181
<i>Kai Zhao, Liting Ruan, Haoran Jiang, Xiaoqiang Zhu, Xianchao Zhang, Dan Zeng</i>	
Cheating Stereo Matching in Full-Scale: Physical Adversarial Attack Against Binocular Depth Estimation in Autonomous Driving.....	13190
<i>Kangqiao Zhao, Shuo Huai, Xurui Song, Jun Luo</i>	
ControlFuse: Instruction-Guided Multi-Granularity Controllable Image Fusion.....	13199
<i>Libo Zhao, Xiaoli Zhang, Zeyu Wang</i>	
Predicting Video Slot Attention Queries from Random Slot-Feature Pairs.....	13208
<i>Rongzhen Zhao, Jian Li, Juho Kannala, Joni Pajarinen</i>	
Adaptive-Smooth LiDAR-Camera Knowledge Distillation with Heterogeneous Fusion for Multi-View 3D Object Detection.....	13217
<i>Rui Zhao, Shuoyao Wang, Xinhua Zheng, Shijian Gao</i>	
Multi-Modal Assistance for Unsupervised Domain Adaptation on Point Cloud 3D Object Detection	13226
<i>Shenao Zhao, Pengpeng Liang, Zhoufan Yang</i>	
ObjectAdv: Object-Level Unrestricted Adversarial Attacks Via Diffusion Models	13235
<i>Shijie Zhao, Zhenyu Liang, Xing Yang, Haoqi Gao, Anjie Peng, Hui Zeng</i>	
KineST: A Kinematics-Guided Spatiotemporal State Space Model for Human Motion Tracking from Sparse Signals.....	13244
<i>Shuting Zhao, Zeyu Xiao, Xinrong Chen</i>	
DialoGen: Towards Dialog Gesture Generation Via Identity-Decoupled Style Guidance in Interactive Diffusion Model	13253
<i>Weiyu Zhao, Chenyang Wang, Liangxiao Hu, Zonglin Li, Wei Yu, Shengping Zhang</i>	
Good Gradients Poison Your Model: Evading Defenses in Federated Learning Via Boundary-Adaptive Perturbation.....	13262
<i>Xiaojie Zhao, Jinqiao Shi, Yi Li, Junmin Huang, Chongru Fan</i>	
Studying Classifier(-Free) Guidance from a Classifier-Centric Perspective	13271
<i>Xiaoming Zhao, Alex Schwing</i>	
GloTok: Global Perspective Tokenizer for Image Reconstruction and Generation.....	13280
<i>Xuan Zhao, Zhongyu Zhang, Yuge Huang, Yuxi Mi, Guodong Mu, Shouhong Ding, Jun Wang, Rizen Guo, Shuigeng Zhou</i>	
OpenScan: A Benchmark for Generalized Open-Vocabulary 3D Scene Understanding	13289
<i>Youjun Zhao, Jiaying Lin, Shuquan Ye, Qianshi Pang, Rynson W. H. Lau</i>	
UDCH: Unsupervised Dynamic Weighted Cluster-Cooperative Hashing for Cross-Modal Retrieval.....	13297
<i>Yuanzhi Zhao, Fan Yang, Yudong Zhao, Xiaoyu Li</i>	
SceneGenesis: 3D Scene Synthesis Via Semantic Structural Priors and Mesh-Guided Video-Geometry Fusion.....	13305
<i>Yueming Zhao, Hongyu Yang, Di Huang</i>	

Tackling Dual-Stage Missing Modalities in Brain Tumor Segmentation Via Robust Modality Reconstruction and Prompt-Guided Modality Adaptation	13314
<i>Yunpeng Zhao, Cheng Chen, Qing You Pang, Yibing Fu, Quanzheng Li, Carol Tang, Beng Ti Ang, Yueming Jin</i>	
Temporal Calibrating and Distilling for Scene-Text Aware Text-Video Retrieval	13323
<i>Zhiqian Zhao, Liang Li, Lei Shen, Xichun Sheng, Yaoqi Sun, Fang Kang, Chenggang Yan</i>	
CogStream: Context-Guided Streaming Video Question Answering	13332
<i>Zicheng Zhao, Kangyu Wang, Shijie Li, Rui Qian, Weiyao Lin, Huabin Liu</i>	
ProxyTTT: Proxy-Driven Test-Time Training for Multi-Modal Re-Identification.....	13342
<i>Aihua Zheng, Zhaojun Liu, Xixi Wan, Chenglong Li, Jin Tang, Yan Yan</i>	
Progressive Multi-Modal Knowledge Distillation for Multi-Spectral Object Re-Identification	13351
<i>Aihua Zheng, Pengyu Li, Zi Wang, Jin Tang</i>	
Semantic-Driven Visual Progressive Refinement for Aerial-Ground Person ReID: A Challenging Large-Scale Benchmark	13360
<i>Aihua Zheng, Hao Xie, Xixi Wan, Zi Wang, Shihao Li, Jin Tang, Bin Luo</i>	
Manipulating the Mind's Eye: A-SAGE, the Attention-Based Attack on ViT Explainability	13369
<i>Boshi Zheng, Yan Li, Jiabin Liu</i>	
Open-World Deepfake Attribution Via Confidence-Aware Asymmetric Learning.....	13378
<i>Haiyang Zheng, Nan Pu, Wenjing Li, Teng Long, Nicu Sebe, Zhun Zhong</i>	
HiFC-GAN: Hierarchical Feature-Constrained GAN for Optical-To-SAR Transfer in SAR Target Classification	13387
<i>Hao Zheng, Meiguang Zheng, Zhigang Hu, Liu Yang, Aikun Xu, Tingxuan Chen, Rongchang Zhao, Boyu Wang</i>	
Distilling Future Temporal Knowledge with Masked Feature Reconstruction for 3D Object Detection	13396
<i>Haowen Zheng, Hu Zhu, Lu Deng, Weihao Gu, Yang Yang, Yanyan Liang</i>	
Hierarchical Dual-Domain Fusion with Frequency-Guided Spatial Modeling for Pan-Sharpening	13405
<i>Huangqimei Zheng, Chengyi Pan, Qian Jiang, Wei Zhou, Xin Jin</i>	
Universal Adversarial Purification with DDIM Metric Loss for Stable Diffusion	13414
<i>Li Zheng, Liangbin Xie, Jiantao Zhou, He Yimin</i>	
E ³ SAM2: Entropy-Aware and Edge-Guided Adaptation of SAM2 for Echocardiography Video Segmentation	13423
<i>Long Zheng, Zhi Li, Weidong Wang, Zhenyu Dai, Shuyun Li</i>	
WinMamba: Multi-Scale Shifted Windows in State Space Model for 3D Object Detection	13432
<i>Longhui Zheng, Qiming Xia, Xiaolu Chen, Zhaoliang Liu, Chenglu Wen</i>	
Physically-Based LiDAR Smoke Simulation for Robust 3D Object Detection	13441
<i>Shijun Zheng, Yu Guo, Weiquan Liu, Yu Zang, Siqi Shen, Ming Cheng, Cheng Wang</i>	
Forecast Then Calibrate: Feature Caching as ODE for Efficient Diffusion Transformers	13449
<i>Shikang Zheng, Liang Feng, Xinyu Wang, Qinming Zhou, Peiliang Cai, Chang Zou, Jiacheng Liu, Yuqi Lin, Junjie Chen, Yue Ma, Linfeng Zhang</i>	
Content-Aware Information Compression and Selection for Whole Slide Image Analysis	13458
<i>Tingting Zheng, Hongxun Yao, Sicheng Zhao, Yi Xiao</i>	
Selective Diffusion Distillation for Real-World High-Scale Image Super-Resolution.....	13467
<i>Wenli Zheng, Huiyuan Fu, Zekai Xu, Xin Wang, Huadong Ma</i>	
Oscillation Inversion: Training-Free Image and Video Enhancement Through Oscillated Latents in Large Flow Models.....	13476
<i>Yan Zheng, Zhenxiao Liang, Xiaoyan Cong, Yi Yang, Lanqing Guo, Yuehao Wang, Peihao Wang, Zhangyang Wang</i>	
GeoX-Bench: Benchmarking Cross-View Geo-Localization and Pose Estimation Capabilities of Large Multimodal Models	13485
<i>Yushuo Zheng, Jiangyong Ying, Huiyu Duan, Chunyi Li, Zicheng Zhang, Jing Liu, Xiaohong Liu, Guangtao Zhai</i>	

Wavefront-Constrained Passive Obscured Object Detection.....	13494
<i>Zhiwen Zheng, Yiwei Ouyang, Zhao Huang, Tao Zhang, Xiaoshuai Zhang, Huiyu Zhou, Wenwen Tang, Shaowei Jiang, Jin Liu, Xingru Huang</i>	
OwlCap: Harmonizing Motion-Detail for Video Captioning Via HMD-270K and Caption Set Equivalence Reward.....	13503
<i>Chunlin Zhong, Qiuxia Hou, Zhangjun Zhou, Yanhao Zhang, Shuang Hao, Haonan Lu, He Tang, Xiang Bai</i>	
SE360: Semantic Edit in 360° Panoramas Via Hierarchical Data Construction.....	13512
<i>Haoyi Zhong, Fang-Lue Zhang, Andrew Chalmers, Taehyun Rhee</i>	
Collaboratively “Copy & Paste” 2D-3D Features for Complex Video-To-Video Motion Editing	13521
<i>Jia-Xing Zhong, Shijie Zhao, Junlin Li, Li Zhang</i>	
CompEvent: Complex-Valued Event-RGB Fusion for Low-Light Video Enhancement and Deblurring	13530
<i>Mingchen Zhong, Xin Lu, Dong Li, Senyan Xu, Ruixuan Jiang, Xueyang Fu, Baocai Yin</i>	
Empowering Sparse-Input Neural Radiance Fields with Dual-Level Semantic Guidance from Dense Novel Views	13539
<i>Yingji Zhong, Kaichen Zhou, Zhihao Li, Lanqing Hong, Zhenguo Li, Dan Xu</i>	
MoReMouse: Monocular Reconstruction of Laboratory Mouse.....	13548
<i>Yuan Zhong, Jingxiang Sun, Zhongbin Zhang, Liang An, Yebin Liu</i>	
TG-Field: Geometry-Aware Radiative Gaussian Fields for Tomographic Reconstruction	13557
<i>Yuxiang Zhong, Jun Wei, Chaoqi Chen, Senyou An, Hui Huang</i>	
Splat-SAP: Feed-Forward Gaussian Splatting for Human-Centered Scene with Scale-Aware Point Map Reconstruction.....	13566
<i>Boyao Zhou, Shunyuan Zheng, Zhanfeng Liao, Zihan Ma, Hanzhang Tu, Boning Liu, Yebin Liu</i>	
M3ashy: Multi-Modal Material Synthesis Via Hyperdiffusion	13575
<i>Chenliang Zhou, Zheyuan Hu, Alejandro Sztrajman, Yancheng Cai, Yaru Liu, Cengiz Oztireli</i>	
Paper Folding Puzzles: Can Multimodal Large Language Models Perform Spatial Reasoning?.....	13584
<i>Dibin Zhou, Yantao Xu, Zongming Huang, Zengwei Yan, Wenhao Liu, Yongwei Miao, Jianfeng Ren, Fuchang Liu</i>	
IdentityStory: Taming Your Identity-Preserving Generator for Human-Centric Story Generation	13593
<i>Donghao Zhou, Jingyu Lin, Guibao Shen, Quande Liu, Jialin Gao, Lihao Liu, Lan Du, Cunjian Chen, Chi-Wing Fu, Xiaowei Hu, Pheng-Ann Heng</i>	
Exploring Position Encoding Mechanism in Diffusion U-Net for Training-Free High-Resolution Image Generation	13602
<i>Feng Zhou, Pu Cao, Yiyang Ma, Lu Yang, Yonghao Dang, Jianqin Yin</i>	
Toward Real-World High-Precision Image Matting and Segmentation	13611
<i>Haipeng Zhou, Zhaohu Xing, Hongqiu Wang, Jun Ma, Ping Li, Lei Zhu</i>	
MACS: Multi-Source Audio-To-Image Generation with Contextual Significance and Semantic Alignment.....	13620
<i>Hao Zhou, Xiaobao Guo, Yuzhe Zhu, Adams Wai-Kin Kong</i>	
Duplex Rewards Optimization for Test-Time Composed Image Retrieval	13629
<i>Haoliang Zhou, Feifei Zhang, Changsheng Xu</i>	
Seeing and Knowing in the Wild: Open-Domain Visual Entity Recognition with Large-Scale Knowledge Graphs Via Contrastive Learning	13638
<i>Hongkuan Zhou, Lavdim Halilaj, Sebastian Monka, Stefan Schmid, Yuqicheng Zhu, Jingcheng Wu, Nadeem Nazer, Steffen Staab</i>	
T-Rex-Omni: Integrating Negative Visual Prompt in Generic Object Detection.....	13647
<i>Jiazhou Zhou, Qing Jiang, Kanghao Chen, Lutao Jiang, Yuanhuiyi Lyu, Ying-Cong Chen, Lei Zhang</i>	
StrokeFusion: Vector Sketch Generation Via Joint Stroke-UDF Encoding and Latent Sequence Diffusion.....	13656
<i>Jin Zhou, Yi Zhou, Hongliang Yang, Pengfei Xu, Hui Huang</i>	

Think Before You Segment: An Object-Aware Reasoning Agent for Referring Audio-Visual Segmentation.....	13665
<i>Jinxing Zhou, Yanghao Zhou, Mingfei Han, Tong Wang, Xiaojun Chang, Hisham Cholakkal, Rao Muhammad Anwer</i>	
CLASP: Cross-Modal Salient Anchor-Based Semantic Propagation for Weakly-Supervised Dense Audio-Visual Event Localization	13674
<i>Jinxing Zhou, Ziheng Zhou, Yanghao Zhou, Yuxin Mao, Zhangling Duan, Dan Guo</i>	
VPHO: Joint Visual-Physical Cue Learning and Aggregation for Hand-Object Pose Estimation.....	13683
<i>Jun Zhou, Chi Xu, Kaifeng Tang, Yuting Ge, Tingrui Guo, Li Cheng</i>	
Preserving Topological and Geometric Embeddings for Point Cloud Recovery.....	13692
<i>Kaiyue Zhou, Zelong Tan, Hongxiao Wang, Ya-Li Li, Shengjin Wang</i>	
Certified L2-Norm Robustness of 3D Point Cloud Recognition in the Frequency Domain.....	13701
<i>Liang Zhou, Qiming Wang, Tianze Chen</i>	
TargetVAU: Multimodal Anomaly-Aware Reasoning for Target Behavior Understanding in Videos.....	13710
<i>Lingru Zhou, Peng Wu, Manqing Zhang, Qingsheng Wang, Guansong Pang, Peng Wang</i>	
Mitigating Entity Hallucinations in 3D Radiology Report Generation Via Dual-Stream Alignment.....	13719
<i>Lingyu Zhou, Yue Yu, Zhang Yi, Xiuyuan Xu</i>	
Hierarchical Prompt Learning for Image- And Text-Based Person Re-Identification.....	13728
<i>Linhan Zhou, Shuang Li, Neng Dong, Yonghang Tai, Yafei Zhang, Huafeng Li</i>	
Thermal-Physics Guided Infrared Image Super-Resolution with Dynamic High-Frequency Amplification.....	13737
<i>Mingxuan Zhou, Yirui Shen, Shuang Li, Jing Geng, Yutang Zhang, Shuigen Wang</i>	
Reasoning Via Implicit Self-Supervised Emergence for Instruction Segmentation	13746
<i>Qing Zhou, Lichang Yang, Yuyu Jia, Junyu Gao, Weiping Ni, Junzheng Wu, Qi Wang</i>	
Bridging Vision and Language for Robust Context-Aware Surgical Point Tracking: The VL-SurgPT Dataset and Benchmark.....	13755
<i>Rulin Zhou, Wenlong He, An Wang, Jianhang Zhang, Xuanhui Zeng, Xi Zhang, Chaowei Zhu, Haijun Hu, Hongliang Ren</i>	
ASSIST-3D: Adapted Scene Synthesis for Class-Agnostic 3D Instance Segmentation.....	13764
<i>Shengchao Zhou, Jiehong Lin, Jiahui Liu, Shizhen Zhao, Chirui Chang, Xiaojuan Qi</i>	
CompTrack: Information Bottleneck-Guided Low-Rank Dynamic Token Compression for Point Cloud Tracking.....	13773
<i>Sifan Zhou, Yichao Cao, Jiahao Nie, Yuqian Fu, Ziyu Zhao, Xiaobo Lu, Shuo Wang</i>	
OpenDriveVLA: Towards End-To-End Autonomous Driving with Large Vision Language Action Model.....	13782
<i>Xingcheng Zhou, Xuyuan Han, Feng Yang, Yunpu Ma, Volker Tresp, Alois Knoll</i>	
ReaSon: Reinforced Causal Search with Information Bottleneck for Video Understanding.....	13791
<i>Yuan Zhou, Litao Hua, Shilong Jin, Wentao Huang, Haoran Duan</i>	
LookFlow: Training-Free and Efficient High-Resolution Image Synthesis Via Dynamic Lookahead Guidance Flow	13800
<i>Yuan Zhou, Yan Zhang, Jianlong Chang, Xin Gu, Ying Wang, Kun Ding, Guangwen Yang, Shiming Xiang</i>	
Beyond Counting: Evaluating Abstract and Emotional Reasoning in Vision-Language Models.....	13809
<i>Yuan Zhou, Yan Zhang, Jianlong Chang, Xin Gu, Ying Wang, Kun Ding, Guangwen Yang, Shiming Xiang</i>	
DragNeXt: Rethinking Drag-Based Image Editing.....	13818
<i>Yuan Zhou, Junbao Zhou, Qingshan Xu, Kesen Zhao, Yuxuan Wang, Hao Fei, Richang Hong, Hanwang Zhang</i>	
Less is More: Vision Representation Compression for Efficient Video Generation with Large Language Models	13826
<i>Yucheng Zhou, Jihai Zhang, Guanjie Chen, Jianbing Shen, Yu Cheng</i>	

Debiased Dual-Invariant Defense for Adversarially Robust Person Re-Identification.....	13835
<i>Yuhang Zhou, Yanxiang Zhao, Zhongyun Hua, Zhipu Liu, Zhaoquan Gu, Qing Liao, Leo Yu Zhang</i>	
Zero-Shot Open-Vocabulary Human Motion Grounding with Test-Time Training.....	13844
<i>Yunjiao Zhou, Xinyan Chen, Junlang Qian, Lihua Xie, Jianfei Yang</i>	
Few-Step Flow for 3D Generation Via Marginal-Data Transport Distillation.....	13853
<i>Zanwei Zhou, Taoran Yi, Jiemin Fang, Chen Yang, Lingxi Xie, Xinggang Wang, Wei Shen, Qi Tian</i>	
Δ t-Mamba3D: A Time-Aware Spatio-Temporal State-Space Model for Breast Cancer Risk Prediction	13862
<i>Zhengbo Zhou, Dooman Arefan, Margarita Louise Zuley, Shandong Wu</i>	
Semantic Guided Part Relation-Aware Network for Point Cloud Completion.....	13871
<i>Zhensheng Zhou, Jianqing Liang, Jiye Liang, Zijin Du, Chenghao Fang</i>	
DICE: Distilling Classifier-Free Guidance into Text Embeddings.....	13880
<i>Zhenyu Zhou, Defang Chen, Can Wang, Chun Chen, Siwei Lyu</i>	
Content Diversity-Guided Ambiguity Mitigation for Open-Set Noisy Label Learning	13889
<i>Zhihao Zhou, Rui Li, Xueying Li</i>	
HiTVideo: Hierarchical Tokenizers for Enhancing Text-To-Video Generation with Autoregressive Large Language Models	13898
<i>Ziqin Zhou, Yifan Yang, Yuqing Yang, Tianyu He, Houwen Peng, Kai Qiu, Qi Dai, Lili Qiu, Chong Luo, Lingqiao Liu</i>	
Hierarchical Schedule Optimization for Fast and Robust Diffusion Model Sampling.....	13907
<i>Aihua Zhu, Rui Su, Qinglin Zhao, Li Feng, Meng Shen, Shibo He</i>	
MedEyes: Learning Dynamic Visual Focus for Medical Progressive Diagnosis	13916
<i>Chunzheng Zhu, Yangfang Lin, Shen Chen, Yijun Wang, Jianxin Lin</i>	
Self-Supervised Representation Learning with Joint Embedding Predictive Architecture for Automotive LiDAR Object Detection.....	13925
<i>Haoran Zhu, Zhenyuan Dong, Kristi Topollai, Beiyao Sha, Anna Ewa Choromanska</i>	
Other Vehicle Trajectories Are Also Needed: A Driving World Model Unifies Ego-Other Vehicle Trajectories in Video Latent Space.....	13934
<i>Jian Zhu, Zhengyu Jia, Tian Gao, Jiaxin Deng, Shidi Li, Lang Zhang, Fu Liu, Peng Jia, Xianpeng Lang</i>	
AnchorDS: Anchoring Dynamic Sources for Semantically Consistent Text-To-3D Generation.....	13943
<i>Jiayin Zhu, Linlin Yang, Yicong Li, Angela Yao</i>	
LENS: Learning to Segment Anything with Unified Reinforced Reasoning	13952
<i>Lianghui Zhu, Bin Ouyang, Yuxuan Zhang, Tianheng Cheng, Rui Hu, Haocheng Shen, Longjin Ran, Xiaoxin Chen, Li Yu, Wenyu Liu, Xinggang Wang</i>	
Class Incremental Medical Image Segmentation Via Prototype-Guided Calibration and Dual-Aligned Distillation	13961
<i>Shengqian Zhu, Chengrong Yu, Qiang Wang, Ying Song, Guangjun Li, Jiafei Wu, Xiaogang Xu, Zhang Yi, Junjie Hu</i>	
Graph-Driven Domain Co-Adaptation for Cross-Domain Image Quality Assessment	13970
<i>Shun Zhu, Xichen Yang, Yan Zhang, Tianshu Wang, Zhongyuan Mao, Tianyin Li, Zhuoyan Sun, Xiaobo Shen</i>	
VTD-CLIP: Video-To-Text Discretization Via Prompting CLIP	13979
<i>Wencheng Zhu, Yuexin Wang, Hongxuan Li, Pengfei Zhu</i>	
SLCFormer: Spectral-Local Context Transformer with Physics-Grounded Flare Synthesis for Nighttime Flare Removal	13988
<i>Xiyu Zhu, Wei Wang, Xin Yuan, Xiao Wang</i>	
CapeNext: Rethinking and Refining Dynamic Support Information for Category-Agnostic Pose Estimation.....	13997
<i>Yu Zhu, Dan Zeng, Shuiwang Li, Qijun Zhao, Qiaomu Shen, Bo Tang</i>	

Pixel-Level Quality Assessment for Oriented Object Detection	14005
<i>Yunhui Zhu, Buliao Huang</i>	
SinBasis Networks: Matrix-Equivalent Feature Extraction for Wave-Like Optical Spectrograms	14014
<i>Yuzhou Zhu, Zheng Zhang, Ruyi Zhang, Liang Zhou</i>	
Training-Free Spatio-Temporal Decoupled Reasoning Video Segmentation with Adaptive Object Memory	14022
<i>Zhengdong Zhu, Jiaqing Fan, Zhixuan Liu, Fanzhang Li</i>	
Q Cache: Visual Attention is Valuable in Less than Half of Decode Layers for Multimodal Large Language Model.....	14031
<i>Jiedong Zhuang, Lu Lu, Ming Dai, Rui Hu, Jian Chen, Qiang Liu, Haoji Hu</i>	
Libra-MIL: Multimodal Prototypes Stereoscopic Infused with Task-Specific Language Priors for Few-Shot Whole Slide Image Classification	14040
<i>Zhenfeng Zhuang, Fangyu Zhou, Liansheng Wang</i>	
Tuning for Two Adversaries: Enhancing the Robustness Against Transfer and Query-Based Attacks Using Hyperparameter Tuning	14049
<i>Pascal Zimmer, Ghassan Karame</i>	
HouseTune: Two-Stage Floorplan Generation with LLM Assistance	14059
<i>Ziyang Zong, Guanying Chen, Zhaohuan Zhan, Fengcheng Yu, Guang Tan</i>	
EchoGen: Cycle-Consistent Learning for Unified Layout-Image Generation and Understanding	14068
<i>Kai Zou, Hongbo Liu, Dian Zheng, Jianxiong Gao, Zhiwei Zhao, Bin Liu</i>	
Appearance Discrepancy-Guided Sequence Hybrid Masking for Robust Scene Text Recognition	14077
<i>Shihao Zou, Wei Wei, Leyang Xu, Kaihe Xu, Wenfeng Xie</i>	
Turbo-VAED: Fast and Stable Transfer of Video-VAEs to Mobile Devices	14086
<i>Ya Zou, Jingfeng Yao, Siyuan Yu, Shuai Zhang, Wenyu Liu, Xinggang Wang</i>	
HATIR: Heat-Aware Diffusion for Turbulent Infrared Video Super-Resolution	14095
<i>Yang Zou, Xingyue Zhu, Kaiqi Han, Jun Ma, Xingyuan Li, Zhiying Jiang, Jinyuan Liu</i>	
Boosting Adversarial Transferability Via Ensemble Non-Attention.....	14104
<i>Yipeng Zou, Qin Liu, Jie Wu, Yu Peng, Guo Chen, Hui Zhou, Guanghui Ye</i>	
Decoupling What to Count and Where to See for Referring Expression Counting	14113
<i>Yuda Zou, Zijian Zhang, Yongchao Xu</i>	
CoordAR: One-Reference 6D Pose Estimation of Novel Objects Via Autoregressive Coordinate Map Generation	14122
<i>Dexin Zuo, Ang Li, Wei Wang, Wenxian Yu, Danping Zou</i>	

Author Index