

**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 4: AAAI Technical Tracks

- Computer Vision I

Singapore
20-27 January 2026

ISBN: 979-8-3313-3378-2

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 I

| | |
|---|------|
| LiteGE: Lightweight Geodesic Embedding for Efficient Geodesics Computation and Non-Isometric Shape Correspondence | 2291 |
| <i>Yohanes Yudhi Adikusuma, Qixing Huang, Ying He</i> | |
| Open-World Object Counting in Videos..... | 2300 |
| <i>Niki Amini-Naieni, Andrew Zisserman</i> | |
| Spatio-Temporal Distortion Aware Omnidirectional Video Super-Resolution | 2309 |
| <i>Hongyu An, Xinfeng Zhang, Shijie Zhao, Li Zhang, Ruiqin Xiong</i> | |
| Enhancing Retrieval-Augmented Large Vision Language Models Via Knowledge Conflict Mitigation | 2318 |
| <i>Wenbin An, Jiahao Nie, Feng Tian, Mingxiang Cai, Yaqiang Wu, Xiaoqin Zhang, Shijian Lu</i> | |
| Cross-Temporal 3D Gaussian Splatting for Sparse-View Guided Scene Update | 2327 |
| <i>Zeyuan An, Yanghang Xiao, Zhiying Leng, Frederick W. B. Li, Xiaohui Liang</i> | |
| Towards Temporal Fusion Beyond the Field of View for Camera-Based Semantic Scene Completion | 2336 |
| <i>Jongseong Bae, Junwoo Ha, Jinnyeong Heo, Yeongin Lee, Ha Young Kim</i> | |
| DogFit: Domain-Guided Fine-Tuning for Efficient Transfer Learning of Diffusion Models..... | 2345 |
| <i>Yara Bahram, Mohammadhadi Shateri, Eric Granger</i> | |
| Learning Compact Latent Space for Representing Neural Signed Distance Functions with High-Fidelity Geometry Details | 2354 |
| <i>Qiang Bai, Bojian Wu, Xi Yang, Zhizhong Han</i> | |
| HyperCOD: The First Challenging Benchmark and Baseline for Hyperspectral Camouflaged Object Detection..... | 2363 |
| <i>Shuyan Bai, Tingfa Xu, Peifu Liu, Yuhao Qiu, Huiyan Bai, Huan Chen, Yanyan Peng, Jianan Li</i> | |
| Plug-And-Play Optimization for 3D Gaussian Splatting Compression: Distribution Regularization, Probabilistic Pruning and Detail Compensation..... | 2372 |
| <i>Tian Bai, Zheng Qiu, Haojie Chen, Ziyang Dai</i> | |
| SDNet: LiDAR Semantic Scene Completion with Sparse-Dense Fusion and Input-Aware Label Refinement | 2381 |
| <i>Tingming Bai, Zhiyu Xiang, Peng Xu, Tianyu Pu, Kai Wang, Eryun Liu</i> | |
| Complex Mathematical Expression Recognition: Benchmark, Large-Scale Dataset and Strong Baseline | 2390 |
| <i>Weikang Bai, Yongkun Du, Yuchen Su, Yazhen Xie, Zhineng Chen</i> | |
| MCIE: Multimodal LLM-Driven Complex Instruction Image Editing with Spatial Guidance..... | 2399 |
| <i>Xuehai Bai, Xiaoling Gu, Akide Liu, Hangjie Yuan, Yifan Zhang, Jack Ma</i> | |

| | |
|---|------|
| Stop Mixing Things Up! BISCUIT Teaches Vision-Language Models to Learn New Concepts from Images on the Spot | 2408 |
| <i>Jiahua Bao, Siyao Cheng, Jiaxing Du, Yuhang Jia, Boyang Niu, Zeming Lang, Changjiang He, Hao Zhang, Jie Liu</i> | |
| TripleFDS: Triple Feature Disentanglement and Synthesis for Scene Text Editing | 2417 |
| <i>Yuchen Bao, Yiting Wang, Wenjian Huang, Haowei Wang, Shen Chen, Taiping Yao, Shouhong Ding, Jianguo Zhang</i> | |
| CPOStream: Collaborating Prediction and Observation for Flicker-Free Streamable Free-Viewpoint Video with 3DGS | 2426 |
| <i>Zhenyu Bao, Qing Li, Jinhan Xie, Kanglin Liu</i> | |
| Text-To-Scene with Large Reasoning Models..... | 2435 |
| <i>Frédéric Berdoz, Luca A Lanzendörfer, Nick Tuninga, Roger Wattenhofer</i> | |
| SSR: Semantic and Spatial Rectification for CLIP-Based Weakly Supervised Segmentation | 2444 |
| <i>Xiuli Bi, Die Xiao, Junchao Fan, Bin Xiao</i> | |
| Foundation-Adaptive Integrated Refinement for Generalized Category Discovery..... | 2453 |
| <i>Yuwei Bian, Shidong Wang, Yazhou Yao, Haofeng Zhang</i> | |
| Robust Pedestrian Detection with Uncertain Modality | 2462 |
| <i>Qian Bie, Xiao Wang, Bin Yang, Zhixi Yu, Jun Chen, Xin Xu</i> | |
| Knowledge-Enhanced Explainable Prompting for Vision-Language Models | 2471 |
| <i>Yequan Bie, Andong Tan, Zhixuan Chen, Zhiyuan Cai, Luyang Luo, Hao Chen</i> | |
| MoBGS: Motion Deblurring Dynamic 3D Gaussian Splatting for Blurry Monocular Video | 2480 |
| <i>Minh-Quan Viet Bui, Jongmin Park, Juan Luis Gonzalez, Jaeho Moon, Jihyong Oh, Munchurl Kim</i> | |
| DOS: Directional Object Separation in Text Embeddings for Multi-Object Image Generation..... | 2490 |
| <i>Dongnam Byun, Jungwon Park, Jungmin Ko, Changin Choi, Wonjong Rhee</i> | |
| Symmetrical Flow Matching: Unified Image Generation, Segmentation, and Classification with Score-Based Generative Models | 2498 |
| <i>Francisco Caetano, Christiaan Viviers, Peter H. N. De With, Fons Van Der Sommen</i> | |
| FastFLUX: Pruning FLUX with Block-Wise Replacement and Sandwich Training | 2507 |
| <i>Fuhan Cai, Yong Guo, Jie Li, Wenbo Li, Jian Chen, Xiangzhong Fang</i> | |
| Seeing in Double: Dual-Granularity BEV Segmentation Via Mamba-Driven Alignment and Polar-Decoupled Experts | 2516 |
| <i>Jiaxin Cai, Rui Lin, Jingze Su, Qi Li, Wenjie Yang, Yuanlong Yu, Wenxi Liu</i> | |
| DriveLiDAR4D: Sequential and Controllable LiDAR Scene Generation for Autonomous Driving | 2525 |
| <i>Kaiwen Cai, Xinze Liu, Xia Zhou, Hengtong Hu, Jie Xiang, Luyao Zhang, Xueyang Zhang, Kun Zhan, Yifei Zhan, Xianpeng Lang</i> | |
| Zooming in on Fakes: A Novel Dataset for Localized AI-Generated Image Detection with Forgery Amplification Approach | 2534 |
| <i>Lvpan Cai, Haowei Wang, Jiayi Ji, Yanshu Zhoumen, Shen Chen, Taiping Yao, Xiaoshuai Sun</i> | |
| SEMC: Structure-Enhanced Mixture-Of-Experts Contrastive Learning for Ultrasound Standard Plane Recognition..... | 2543 |
| <i>Qing Cai, Guihaoyan, Fan Zhang, Cheng Zhang, Zhi Liu</i> | |

| | |
|--|------|
| Mem4D: Decoupling Static and Dynamic Memory for Dynamic Scene Reconstruction | 2552 |
| <i>Xudong Cai, Shuo Wang, Peng Wang, Yongcai Wang, Zhaoxin Fan, Wanting Li, Tianbao Zhang, Jianrong Tao, Yeying Jin, Deying Li</i> | |
| Split-Layer: Enhancing Implicit Neural Representation by Maximizing the Dimensionality of Feature Space | 2561 |
| <i>Zhicheng Cai, Hao Zhu, Linsen Chen, Qiu Shen, Xun Cao</i> | |
| FastDriveVLA: Efficient End-To-End Driving Via Plug-And-Play Reconstruction-Based Token Pruning | 2571 |
| <i>Jiajun Cao, Qizhe Zhang, Peidong Jia, Xuhui Zhao, Bo Lan, Xiaoan Zhang, Lizhuo , Xiaobao Wei, Sixiang Chen, Liyun Li, Xianming Liu, Ming Lu, Yang Wang, Shanghang Zhang</i> | |
| Learning Depth from Past Selves: Self-Evolution Contrast for Robust Depth Estimation | 2580 |
| <i>Jing Cao, Kui Jiang, Shenyi Li, Xiaocheng Feng, Yong Huang</i> | |
| UQ-Bench: A Benchmark for Evaluating Multimodal LLMs on Underwater Image Quality Assessment | 2589 |
| <i>Jingchao Cao, Guo An, Feng Gao, Ke Gu, Yutao Liu</i> | |
| RelaCtrl: Relevance-Guided Efficient Control for Diffusion Transformers | 2598 |
| <i>Ke Cao, Jing Wang, Ao Ma, Jiasong Feng, Xuanhua He, Run Ling, Haowei Liu, Jian Lu, Wei Feng, Haozhe Wang, Hongjuan Pei, Yihua Shao, Zhanjie Zhang, Jie Zhang</i> | |
| VQAThinker: Exploring Generalizable and Explainable Video Quality Assessment Via Reinforcement Learning | 2607 |
| <i>Linhan Cao, Wei Sun, Weixia Zhang, Xiangyang Zhu, Jun Jia, Kaiwei Zhang, Dandan Zhu, Guangtao Zhai, Xiongkuo Min</i> | |
| Video SimpleQA: Towards Factuality Evaluation in Large Video Language Models..... | 2616 |
| <i>Meng Cao, Pengfei Hu, Yingyao Wang, Jihao Gu, Haoran Tang, Haoze Zhao, Chen Wang, Jiahua Dong, Wangbo Yu, Ge Zhang, Xiang Li, Ian Reid, Xiaodan Liang</i> | |
| Latent Knowledge-Guided Video Diffusion for Scientific Phenomena Generation from a Single Initial Frame | 2625 |
| <i>Qinglong Cao, Xirui Li, Ding Wang, Chao Ma, Yuntian Chen, Xiaokang Yang</i> | |
| EAGLE: Episodic Appearance- And Geometry-Aware Memory for Unified 2D-3D Visual Query Localization in Egocentric Vision | 2634 |
| <i>Yifei Cao, Yu Liu, Guolong Wang, Zhu Liu, Kai Wang, Xianjie Zhang, Jizhe Yu, Xun Tu</i> | |
| Robust Noise Modeling for Spike Camera Via Time-Interval Quantification and Spike-DSLR Multimodal Dataset in Low-Light Imaging..... | 2643 |
| <i>Yue Cao, Sizhao Li, Liguang Zhang</i> | |
| Automatic Translational Correction of Multi-View Coronary Angiography Based on Auto-Annotation Data Generation | 2652 |
| <i>Yue Cao, Zhuo Zhang, Shuai Xiao, Jialin Li, Guipeng Lan, Jiabao Wen, Jiachen Yang</i> | |
| Audio-Assisted Face Video Restoration with Temporal and Identity Complementary Learning..... | 2661 |
| <i>Yuqin Cao, Yixuan Gao, Wei Sun, Xiaohong Liu, Yulun Zhang, Xiongkuo Min</i> | |
| Multivariate Diffusion Transformer with Decoupled Attention for High-Fidelity Mask-Text Collaborative Facial Generation | 2670 |
| <i>Yushe Cao, Dianxi Shi, Xing Fu, Xuechao Zou, Haikuo Peng, Xueqi Li, Chun Yu, Junliang Xing</i> | |

| | |
|--|------|
| Maniflat3D: Learning 3D Geometry Through Planar Representations from Multi-Layer Unwrapping | 2680 |
| <i>Zijian Cao, Dayou Zhang, Zeyuan Liu, Zhicheng Liang, Fangxin Wang</i> | |
| QuantVSR: Low-Bit Post-Training Quantization for Real-World Video Super-Resolution..... | 2689 |
| <i>Bowen Chai, Zheng Chen, Libo Zhu, Wenbo Li, Yong Guo, Yulun Zhang</i> | |
| AbductiveMLLM: Boosting Visual Abductive Reasoning Within MLLMs..... | 2698 |
| <i>Boyu Chang, Qi Wang, Xi Guo, Zhixiong Nan, Yazhou Yao, Tianfei Zhou</i> | |
| Reconstruction Using the Invisible: Intuition from NIR and Metadata for Enhanced 3D Gaussian Splatting | 2707 |
| <i>Gyusam Chang, Tuan-Anh Vu, Vivek Alumootil, Harris Song, Deanna Pham, Sangpil Kim, M. Khalid Jawed</i> | |
| MeshSplat: Generalizable Sparse-View Surface Reconstruction Via Gaussian Splatting | 2716 |
| <i>Hanzhi Chang, Ruijie Zhu, Wenjie Chang, Mulin Yu, Yanzhe Liang, Jiahao Lu, Zhuoyuan Li, Tianzhu Zhang</i> | |
| MambaOVSR: Multiscale Fusion with Global Motion Modeling for Chinese Opera Video Super-Resolution..... | 2725 |
| <i>Hua Chang, Xin Xu, Wei Liu, Wei Wang, Xin Yuan, Kui Jiang</i> | |
| Escaping the CAM Shadow: Uncertainty-Guided Reliable Learning for Weakly Supervised Semantic Segmentation | 2734 |
| <i>Luyao Chang, Leiting Chen, Chen Yang, Chuan Zhou</i> | |
| BulletTime4D: Towards High Spatio-Temporal Resolution Dynamic Scene Rendering Via Spike-Guided Stereo Vision..... | 2743 |
| <i>Yiqian Chang, Haoran Xu, Qinghong Ye, Jianing Li, Xuan Wang, Wei Zhang, Peixi Peng</i> | |
| PerTouch: VLM-Driven Agent for Personalized and Semantic Image Retouching | 2752 |
| <i>Zewei Chang, Zheng-Peng Duan, Jianxing Zhang, Chun-Le Guo, Siyu Liu, Hyungju Chun, Hyunhee Park, Zikun Liu, Chongyi Li</i> | |
| Mitigating Perception Bias: A Training-Free Approach to Enhance LMM for Image Quality Assessment | 2760 |
| <i>Baoliang Chen, Siyi Pan, Dongxu Wu, Liang Xie, Xiangjie Sui, Lingyu Zhu, Hanwei Zhu</i> | |
| Style4D-Bench: A Benchmark Suite for 4D Stylization..... | 2769 |
| <i>Beiqi Chen, Shuai Shao, Haitang Feng, Jianhuang Lai, Jianlou Si, Guangcong Wang</i> | |
| Towards Ultrasound-Based Reliable Disease Diagnosis Using Causal Inference | 2778 |
| <i>Bolei Chen, Jiayu Kang, Haonan Yang, Ping Zhong, Yixiong Liang, Rui Fan, Jianxin Wang</i> | |
| Perspective from a Broader Context: Can Room Style Knowledge Help Visual Floorplan Localization? | 2787 |
| <i>Bolei Chen, Shengsheng Yan, Yongzheng Cui, Jiayu Kang, Ping Zhong, Jianxin Wang</i> | |
| Vision-Only Gaussian Splatting for Collaborative Semantic Occupancy Prediction | 2796 |
| <i>Cheng Chen, Hao Huang, Saurabh Bagchi</i> | |
| AerialMind: Towards Referring Multi-Object Tracking in UAV Scenarios | 2805 |
| <i>Chenglizhao Chen, Shaofeng Liang, Runwei Guan, Xiaolou Sun, Haocheng Zhao, Haiyun Jiang, Tao Huang, Henghui Ding, Qing-Long Han</i> | |

| | |
|--|------|
| Action-And-Object Aware Alignment for Partially Relevant Video Retrieval | 2814 |
| <i>Chuanshen Chen, Kai Zhou, Zhiquan Wen, Zeng You, Yirui Li, Tianhang Xiang, Mingkui Tan</i> | |
| UniABG: Unified Adversarial View Bridging and Graph Correspondence for Unsupervised Cross-View Geo-Localization..... | 2823 |
| <i>Cuiqun Chen, Qi Chen, Bin Yang, Xingyi Zhang</i> | |
| Vision-Language Incremental Learning with Dual Class-Individual Memory..... | 2832 |
| <i>Fuhai Chen, Feng Zhang, Xiaoguang Ma, Yiyi Zhou, Jiarong Liu, Xuri Ge</i> | |
| CoRA: A Collaborative Robust Architecture with Hybrid Fusion for Efficient Perception | 2841 |
| <i>Gong Chen, Chaokun Zhang, Pengcheng Lv, Xiaohui Xie</i> | |
| VPSentry: Semi-Supervised Video Polyp Segmentation Via Sentry-Guided Long-Term Prototype Fusion with Correlation Dynamic Propagation | 2850 |
| <i>Guilian Chen, Xiaoling Luo, Huisi Wu, Jing Qin</i> | |
| Sortblock: Similarity-Aware Feature Reuse for Diffusion Model | 2859 |
| <i>Hanqi Chen, Xu Zhang, Xiaoliu Guan, Lielin Jiang, Guanzhong Wang, Zeyu Chen, Yi Liu</i> | |
| HiFi-Mamba: Dual-Stream γ -Laplacian Enhanced Mamba for High-Fidelity MRI Reconstruction..... | 2868 |
| <i>Hongli Chen, Pengcheng Fang, Yuxia Chen, Yingxuan Ren, Jing Hao, Fangfang Tang, Xiaohao Cai, Shanshan Shan, Feng Liu</i> | |
| CareCom: Generative Image Composition with Calibrated Reference Features..... | 2877 |
| <i>Jiaxuan Chen, Bo Zhang, Qingdong He, Jinlong Peng, Li Niu</i> | |
| LoGoSeg: Integrating Local and Global Features for Open-Vocabulary Semantic Segmentation | 2886 |
| <i>Junyong Chen, Xiangbo Lv, Zhiqiang Kou, Xingdong Sheng, Ning Xu, Yiguo Qiao</i> | |
| Intra-Image Mining and Symmetric Maximum Concept Matching for Few Shot Out-Of-Distribution Detection | 2895 |
| <i>Kaixiang Chen, Pengfei Fang, Hui Xue</i> | |
| EvDiff3D: Event-Aware Diffusion Repair for High-Fidelity Event-Based 3D Reconstruction | 2904 |
| <i>Kanghao Chen, Zixin Zhang, Hangyu Li, Lin Wang, Zeyu Wang</i> | |
| MosaicDoc: A Large-Scale Bilingual Benchmark for Visually Rich Document Understanding..... | 2913 |
| <i>Ketong Chen, Yuhao Chen, Yang Xue</i> | |
| CoMA-SLAM: Collaborative Multi-Agent Gaussian SLAM with Geometric Consistency | 2922 |
| <i>Lin Chen, Yongxin Su, Jvboxi Wang, Pengcheng Han, Zhenyu Xia, Shuhui Bu, Kun Li, Boni Hu, Shengqi Meng, Guangming Wang</i> | |
| StegaVAR: Privacy-Preserving Video Action Recognition Via Steganographic Domain Analysis..... | 2930 |
| <i>Lixin Chen, Chaomeng Chen, Jiale Zhou, Zhijian Wu, Xun Lin</i> | |
| Human-Centric Video Generation Via Collaborative Multi-Modal Conditioning..... | 2939 |
| <i>Liyang Chen, Tianxiang Ma, Jiawei Liu, Bingchuan Li, Zhuowei Chen, Lijie Liu, Xu He, Gen Li, Qian He, Zhiyong Wu</i> | |
| Fast Multi-View Consistent 3D Editing with Video Priors..... | 2948 |
| <i>Liyi Chen, Ruihuang Li, Guowen Zhang, Pengfei Wang, Lei Zhang</i> | |
| Sparse-VDiT: Unleashing the Power of Sparse Attention to Accelerate Video Diffusion Transformers..... | 2957 |
| <i>Pengtao Chen, Xianfang Zeng, Maosen Zhao, Mingzhu Shen, Wei Cheng, Gang Yu, Tao Chen</i> | |

| | |
|--|------|
| 3D-DRES: Detailed 3D Referring Expression Segmentation | 2966 |
| <i>Qi Chen, Changli Wu, Jiayi Ji, Yiwei Ma, Liujuan Cao</i> | |
| LiDAR-GS++: Improving LiDAR Gaussian Reconstruction Via Diffusion Priors..... | 2975 |
| <i>Qifeng Chen, Jiarun Liu, Rengan Xie, Tao Tang, Sicong Du, Yiru Zhao, Yuchi Huo, Sheng Yang</i> | |
| VILTA: A VLM-In-The-Loop Adversary for Enhancing Driving Policy Robustness | 2984 |
| <i>Qimao Chen, Fang Li, Shaoqing Xu, Zhiyi Lai, Zixun Xie, Yuechen Luo, Shengyin Jiang, Hanbing Li, Long Chen, Bing Wang, Yi Zhang, Zhi-Xin Yang</i> | |
| FreeGaussian: Annotation-Free Control of Articulated Objects Via 3D Gaussian Splats with Flow Derivatives..... | 2993 |
| <i>Qizhi Chen, Delin Qu, Junli Liu, Yiwen Tang, Haoming Song, Dong Wang, Yuan Yuan, Bin Zhao</i> | |
| DeNAS-ViT: Data Efficient NAS-Optimized Vision Transformer for Ultrasound Image Segmentation | 3002 |
| <i>Renqi Chen, Xinzhe Zheng, Haoyang Su, Kehan Wu</i> | |
| Domain-Auxiliary Infrared Moving Small Target Detection by Learning to Overlook Domain Discrepancy | 3011 |
| <i>Shengjia Chen, Luping Ji, Shuang Peng, Sicheng Zhu, Mao Ye</i> | |
| Unleashing Semantic and Geometric Priors for 3D Scene Completion | 3020 |
| <i>Shiyuan Chen, Wei Sui, Bohao Zhang, Zeyd Boukhers, John See, Cong Yang</i> | |
| Exploring the Potentials of Spiking Neural Networks for Image Deraining | 3029 |
| <i>Shuang Chen, Tom Krajnik, Farshad Arvin, Amir Atapour-Abarghouei</i> | |
| Spectral Property-Driven Data Augmentation for Hyperspectral Single-Source Domain Generalization | 3038 |
| <i>Taiqin Chen, Yifeng Wang, Xiaochen Feng, Zhilin Zhu, Hao Sha, Yingjian Li, Yongbing Zhang</i> | |
| EndoIR: Degradation-Agnostic All-In-One Endoscopic Image Restoration Via Noise-Aware Routing Diffusion..... | 3047 |
| <i>Tong Chen, Xinyu Ma, Long Bai, Wenyang Wang, Yue Sun, Luping Zhou</i> | |
| IGIANet: Illumination Guided Implicit Alignment Network for Infrared-Visible UAV Detection | 3056 |
| <i>Xiangqi Chen, Dawei Zhang, Li Zhao, Chengzhuan Yang, Zhongyu Chen, Jungang Lou, Zhonglong Zheng, Sang-Woon Jeon, Hua Wang</i> | |
| Unsupervised Multi-View Visual Anomaly Detection Via Progressive Homography-Guided Alignment..... | 3065 |
| <i>Xintao Chen, Xiaohao Xu, Bozhong Zheng, Yun Liu, Yingna Wu</i> | |
| Flowing Backwards: Improving Normalizing Flows Via Reverse Representation Alignment..... | 3074 |
| <i>Yang Chen, Xiaowei Xu, Shuai Wang, Chenhui Zhu, Ruxue Wen, Xubin Li, Tiezheng Ge, Limin Wang</i> | |
| SAM2-OV: A Novel Detection-Only Tuning Paradigm for Open-Vocabulary Multi-Object Tracking..... | 3083 |
| <i>Yangkai Chen, Qiangqiang Wu, Guangyao Li, Junlong Gao, Guanglin Niu, Hanzi Wang</i> | |
| Multimodal Gaussian Mixture Variational Autoencoder with Consistency Regularizations..... | 3092 |
| <i>Yarui Chen, Lehan Hong, Jianlin Shao, Jianning Yang, Tingting Zhao, Yun Liao, Yancui Shi</i> | |

| | |
|---|------|
| ProPL: Universal Semi-Supervised Ultrasound Image Segmentation Via Prompt-Guided Pseudo-Labeling..... | 3101 |
| <i>Yaxiong Chen, Qicong Wang, Chunlei Li, Jingliang Hu, Yilei Shi, Shengwu Xiong, Xiao Xiang Zhu, Lichao Mou</i> | |
| PCGS: Progressive Compression of 3D Gaussian Splatting | 3111 |
| <i>Yihang Chen, Mengyao Li, Qianyi Wu, Weiyao Lin, Mehrtash Harandi, Jianfei Cai</i> | |
| StreamKV: Streaming Video Question-Answering with Segment-Based KV Cache Retrieval and Compression..... | 3120 |
| <i>Yilong Chen, Xiang Bai, Zhibin Wang, Chengyu Bai, Yuhan Dai, Ming Lu</i> | |
| ContextFlow: Training-Free Video Object Editing Via Adaptive Context Enrichment..... | 3129 |
| <i>Yiyang Chen, Xuanhua He, Xiujun Ma, Jack Ma</i> | |
| RefAdGen: High-Fidelity Advertising Image Generation..... | 3138 |
| <i>Yiyun Chen, Weikai Yang</i> | |
| HybridDLA: Hybrid Generation for Document Layout Analysis | 3147 |
| <i>Yufan Chen, Omar Moured, Ruiping Liu, Junwei Zheng, Kunyu Peng, Jiaming Zhang, Rainer Stiefelhagen</i> | |
| Revisiting Network Inertia: Dynamic Inertia Inhibition Coupled Multidimensional Periodicity for Infrared and Visible Image Fusion | 3156 |
| <i>Yufeng Chen, Yuan Sun, Hao Pan, Xujian Zhao, Jian Dai, Zhenwen Ren, Xingfeng Li</i> | |
| O-DisCo-Edit: Object Distortion Control for Unified Realistic Video Editing | 3165 |
| <i>Yuqing Chen, Junjie Wang, Lin Liu, Ruihang Chu, Xiaopeng Zhang, Qi Tian, Yujiu Yang</i> | |
| LAMIC: Layout-Aware Multi-Image Composition Via Scalability of Multimodal Diffusion Transformer | 3174 |
| <i>Yuzhuo Chen, Zehua Ma, Jianhua Wang, Kai Kang, Shunyu Yao, Weiming Zhang</i> | |
| RAPTOR: Real-Time High-Resolution UAV Video Prediction with Efficient Video Attention | 3183 |
| <i>Zhan Chen, Zile Guo, Enze Zhu, Peirong Zhang, Xiaoxuan Liu, Lei Wang, Yidan Zhang</i> | |

Author Index