

38th AAAI Conference on Artificial Intelligence (AAAI-24), 36th Conference on Innovative Applications of Artificial Intelligence (IAAI-24), 14th Symposium on Educational Advances in Artificial Intelligence (EAAI-24)

Volume 5: AAAI Technical Tracks

- Computer Vision IV

Vancouver, Canada
20 – 27 February 2024

Part 1 of 2

ISBN: 979-8-3313-0087-6

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© (2024) by Association for the Advancement of Artificial Intelligence
All rights reserved.

Printed with permission by Curran Associates, Inc. (2024)

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 IV

DiffusionTrack: Diffusion Model for Multi-Object Tracking.....	3991
<i>Run Luo, Zikai Song, Lintao Ma, Jinlin Wei, Wei Yang, Min Yang</i>	
Devignet: High-Resolution Vignetting Removal via a Dual Aggregated Fusion Transformer with Adaptive Channel Expansion	4000
<i>Shenghong Luo, Xuhan Chen, Weiwen Chen, Zinuo Li, Shuqiang Wang, Chi-Man Pun</i>	
AdaFormer: Efficient Transformer with Adaptive Token Sparsification for Image Super-resolution.....	4009
<i>Xiaotong Luo, Zekun Ai, Qiuyuan Liang, Ding Liu, Yuan Xie, Yanyun Qu, Yun Fu</i>	
SkipDiff: Adaptive Skip Diffusion Model for High-Fidelity Perceptual Image Super-resolution.....	4017
<i>Xiaotong Luo, Yuan Xie, Yanyun Qu, Yun Fu</i>	
Modeling Continuous Motion for 3D Point Cloud Object Tracking	4026
<i>Zhipeng Luo, Gongjie Zhang, Changqing Zhou, Zhonghua Wu, Qingyi Tao, Lewei Lu, Shijian Lu</i>	
SGFormer: Semantic Graph Transformer for Point Cloud-Based 3D Scene Graph Generation	4035
<i>Changsheng Lv, Mengshi Qi, Xia Li, Zhengyuan Yang, Huadong Ma</i>	
Privileged Prior Information Distillation for Image Matting.....	4044
<i>Cheng Lyu, Jiake Xie, Bo Xu, Cheng Lu, Han Huang, Xin Huang, Ming Wu, Chuang Zhang, Yong Tang</i>	
FedST: Federated Style Transfer Learning for Non-IID Image Segmentation.....	4053
<i>Boyuan Ma, Xiang Yin, Jing Tan, Yongfeng Chen, Haiyou Huang, Hao Wang, Weihua Xue, Xiaojuan Ban</i>	
SlowTrack: Increasing the Latency of Camera-Based Perception in Autonomous Driving Using Adversarial Examples.....	4062
<i>Chen Ma, Ningfei Wang, Qi Alfred Chen, Chao Shen</i>	
Uncertainty-Aware GAN for Single Image Super Resolution.....	4071
<i>Chenxi Ma</i>	
Stitching Segments and Sentences towards Generalization in Video-Text Pre-training.....	4080
<i>Fan Ma, Xiaojie Jin, Heng Wang, Jingjia Huang, Linchao Zhu, Yi Yang</i>	
Image Captioning with Multi-Context Synthetic Data	4089
<i>Feipeng Ma, Yizhou Zhou, Fengyun Rao, Yueyi Zhang, Xiaoyan Sun</i>	
Directed Diffusion: Direct Control of Object Placement through Attention Guidance	4098
<i>Wan-Duo Kurt Ma, Avisek Lahiri, J. P. Lewis, Thomas Leung, W. Bastiaan Kleijn</i>	
Unifying Visual and Vision-Language Tracking via Contrastive Learning	4107
<i>Yinchao Ma, Yuyang Tang, Wenfei Yang, Tianzhu Zhang, Jinpeng Zhang, Mengxue Kang</i>	

Follow Your Pose: Pose-Guided Text-to-Video Generation Using Pose-Free Videos	4117
<i>Yue Ma, Yingqing He, Xiaodong Cun, Xintao Wang, Siran Chen, Xiu Li, Qifeng Chen</i>	
Let All Be Whitened: Multi-Teacher Distillation for Efficient Visual Retrieval	4126
<i>Zhe Ma, Jianfeng Dong, Shouling Ji, Zhenguang Liu, Xuhong Zhang, Zonghui Wang, Sifeng He, Feng Qian, Xiaobo Zhang, Lei Yang</i>	
Cross-Layer and Cross-Sample Feature Optimization Network for Few-Shot Fine-Grained Image Classification	4136
<i>Zhen-Xiang Ma, Zhen-Duo Chen, Li-Jun Zhao, Zi-Chao Zhang, Xin Luo, Xin-Shun Xu</i>	
LMD: Faster Image Reconstruction with Latent Masking Diffusion	4145
<i>Zhiyuan Ma, Zhihuan Yu, Jianjun Li, Bowen Zhou</i>	
AdapEdit: Spatio-Temporal Guided Adaptive Editing Algorithm for Text-Based Continuity-Sensitive Image Editing	4154
<i>Zhiyuan Ma, Guoli Jia, Bowen Zhou</i>	
Pay Attention to Target: Relation-Aware Temporal Consistency for Domain Adaptive Video Semantic Segmentation	4162
<i>Huayu Mai, Rui Sun, Yuan Wang, Tianzhu Zhang, Feng Wu</i>	
Improving Automatic VQA Evaluation Using Large Language Models	4171
<i>Oscar Mañas, Benno Krojer, Aishwarya Agrawal</i>	
Inconsistency-Based Data-Centric Active Open-Set Annotation	4180
<i>Ruiyu Mao, Ouyang Xu, Yunhui Guo</i>	
Progressive High-Frequency Reconstruction for Pan-Sharpener with Implicit Neural Representation	4189
<i>Ge Meng, Jingjia Huang, Yingying Wang, Zhenqi Fu, Xinghao Ding, Yue Huang</i>	
NaMa: Neighbor-Aware Multi-Modal Adaptive Learning for Prostate Tumor Segmentation on Anisotropic MR Images	4198
<i>Runqi Meng, Xiao Zhang, Shijie Huang, Yuning Gu, Guiqin Liu, Guangyu Wu, Nizhuan Wang, Kaicong Sun, Dinggang Shen</i>	
ConVQG: Contrastive Visual Question Generation with Multimodal Guidance	4207
<i>Li Mi, Syrielle Montariol, Javiera Castillo Navarro, Xianjie Dai, Antoine Bosselut, Devis Tuia</i>	
Out-of-Distribution Detection in Long-Tailed Recognition with Calibrated Outlier Class Learning	4216
<i>Wenjun Miao, Guansong Pang, Xiao Bai, Tianqi Li, Jin Zheng</i>	
BCLNet: Bilateral Consensus Learning for Two-View Correspondence Pruning	4225
<i>Xiangyang Miao, Guobao Xiao, Shiping Wang, Jun Yu</i>	
Understanding the Role of the Projector in Knowledge Distillation	4233
<i>Roy Miles, Krystian Mikolajczyk</i>	
Robust Blind Text Image Deblurring via Maximum Consensus Framework	4242
<i>Zijian Min, Gundu Mohamed Hassan, Geun-Sik Jo</i>	
Knowledge Guided Semi-supervised Learning for Quality Assessment of User Generated Videos	4251
<i>Shankhanil Mitra, Rajiv Soundararajan</i>	
Bridging the Gap between 2D and 3D Visual Question Answering: A Fusion Approach for 3D VQA	4261
<i>Wentao Mo, Yang Liu</i>	

Augmented Commonsense Knowledge for Remote Object Grounding.....	4269
<i>Bahram Mohammadi, Yicong Hong, Yuankai Qi, Qi Wu, Shirui Pan, Javen Qinfeng Shi</i>	
Recurrent Partial Kernel Network for Efficient Optical Flow Estimation.....	4278
<i>Henrique Morimitsu, Xiaobin Zhu, Xiangyang Ji, Xu-Cheng Yin</i>	
TETRIS: Towards Exploring the Robustness of Interactive Segmentation.....	4287
<i>Andrey Moskalenko, Vlad Shakhuro, Anna Vorontsova, Anton Konushin, Anton Antonov, Alexander Krapukhin, Denis Shepelev, Konstantin Soshin</i>	
T2I-Adapter: Learning Adapters to Dig Out More Controllable Ability for Text-to-Image Diffusion Models.....	4296
<i>Chong Mou, Xintao Wang, Liangbin Xie, Yanze Wu, Jian Zhang, Zhongang Qi, Ying Shan</i>	
Semi-supervised Open-World Object Detection.....	4305
<i>Sahal Shaji Mullappilly, Abhishek Singh Gehlot, Rao Muhammad Anwer, Fahad Shahbaz Khan, Hisham Cholakkal</i>	
Adversarial Attacks on the Interpretation of Neuron Activation Maximization.....	4315
<i>Geraldin Nanfack, Alexander Fulleringer, Jonathan Marty, Michael Eickenberg, Eugene Belilovsky</i>	
ColNeRF: Collaboration for Generalizable Sparse Input Neural Radiance Field.....	4325
<i>Zhangkai Ni, Peiqi Yang, Wenhan Yang, Hanli Wang, Lin Ma, Sam Kwong</i>	
Wavelet-Driven Spatiotemporal Predictive Learning: Bridging Frequency and Time Variations.....	4334
<i>Xuesong Nie, Yunfeng Yan, Siyuan Li, Cheng Tan, Xi Chen, Haoyuan Jin, Zhihang Zhu, Stan Z. Li, Donglian Qi</i>	
Painterly Image Harmonization by Learning from Painterly Objects.....	4343
<i>Li Niu, Junyan Cao, Yan Hong, Liqing Zhang</i>	
Progressive Painterly Image Harmonization from Low-Level Styles to High-Level Styles.....	4352
<i>Li Niu, Yan Hong, Junyan Cao, Liqing Zhang</i>	
Domain Generalizable Person Search Using Unreal Dataset.....	4361
<i>Minyoung Oh, Duhyun Kim, Jae-Young Sim</i>	
OctOcc: High-Resolution 3D Occupancy Prediction with Octree.....	4369
<i>Wenzhe Ouyang, Xiaolin Song, Bailan Feng, Zenglin Xu</i>	
NeSyFOLD: A Framework for Interpretable Image Classification.....	4378
<i>Parth Padalkar, Huaduo Wang, Gopal Gupta</i>	
Semi-Supervised Blind Image Quality Assessment through Knowledge Distillation and Incremental Learning.....	4388
<i>Wensheng Pan, Timin Gao, Yan Zhang, Xiawu Zheng, Yunhang Shen, Ke Li, Runze Hu, Yutao Liu, Pingyang Dai</i>	
Less Is More: Label Recommendation for Weakly Supervised Point Cloud Semantic Segmentation.....	4397
<i>Zhiyi Pan, Nan Zhang, Wei Gao, Shan Liu, Ge Li</i>	
patchDPCC: A Patchwise Deep Compression Framework for Dynamic Point Clouds.....	4406
<i>Zirui Pan, Mengbai Xiao, Xu Han, Dongxiao Yu, Guanghui Zhang, Yao Liu</i>	

LISR: Learning Linear 3D Implicit Surface Representation Using Compactly Supported Radial Basis Functions.....	4415
<i>Atharva Pandey, Vishal Yadav, Rajendra Nagar, Santanu Chaudhury</i>	
RadarMOSEVE: A Spatial-Temporal Transformer Network for Radar-Only Moving Object Segmentation and Ego-Velocity Estimation	4424
<i>Changsong Pang, Xieyuanli Chen, Yimin Liu, Huimin Lu, Yuwei Cheng</i>	
NeBLA: Neural Beer-Lambert for 3D Reconstruction of Oral Structures from Panoramic Radiographs.....	4433
<i>Sihwa Park, Seongjun Kim, Doeyoung Kwon, Yohan Jang, In-Seok Song, Seung Jun Baek</i>	
Task-Disruptive Background Suppression for Few-Shot Segmentation.....	4442
<i>Suho Park, SuBeen Lee, Sangeek Hyun, Hyun Seok Seong, Jae-Pil Heo</i>	
SA ² VP: Spatially Aligned-and-Adapted Visual Prompt	4450
<i>Wenjie Pei, Tongqi Xia, Fanglin Chen, Jinsong Li, Jiandong Tian, Guangming Lu</i>	
ConditionVideo: Training-Free Condition-Guided Video Generation.....	4459
<i>Bo Peng, Xinyuan Chen, Yaohui Wang, Chaochao Lu, Yu Qiao</i>	
ViTEraser: Harnessing the Power of Vision Transformers for Scene Text Removal with SegMIM Pretraining	4468
<i>Dezhi Peng, Chongyu Liu, Yuliang Liu, Lianwen Jin</i>	
FRIH: Fine-Grained Region-Aware Image Harmonization.....	4478
<i>Jinlong Peng, Zekun Luo, Liang Liu, Boshen Zhang</i>	
Navigating Open Set Scenarios for Skeleton-Based Action Recognition.....	4487
<i>Kunyu Peng, Cheng Yin, Junwei Zheng, Ruiping Liu, David Schneider, Jiaming Zhang, Kailun Yang, M. Saquib Sarfraz, Rainer Stiefelhagen, Alina Roitberg</i>	
LaneGraph2Seq: Lane Topology Extraction with Language Model via Vertex-Edge Encoding and Connectivity Enhancement.....	4497
<i>Renyuan Peng, Xinyue Cai, Hang Xu, Jiachen Lu, Feng Wen, Wei Zhang, Li Zhang</i>	
Data Adaptive Traceback for Vision-Language Foundation Models in Image Classification	4506
<i>Wenshuo Peng, Kaipeng Zhang, Yue Yang, Hao Zhang, Yu Qiao</i>	
SAM-PARSER: Fine-Tuning SAM Efficiently by Parameter Space Reconstruction	4515
<i>Zelin Peng, Zhengqin Xu, Zhilin Zeng, Xiaokang Yang, Wei Shen</i>	
Relational Distant Supervision for Image Captioning without Image-Text Pairs.....	4524
<i>Yayun Qi, Wentian Zhao, Xinxiao Wu</i>	
Bias-Conflict Sample Synthesis and Adversarial Removal Debias Strategy for Temporal Sentence Grounding in Video	4533
<i>Zhaobo Qi, Yibo Yuan, Xiaowen Ruan, Shuhui Wang, Weigang Zhang, Qingming Huang</i>	
NuScenes-QA: A Multi-Modal Visual Question Answering Benchmark for Autonomous Driving Scenario.....	4542
<i>Tianwen Qian, Jingjing Chen, Linhai Zhuo, Yang Jiao, Yu-Gang Jiang</i>	

PART 2

X-RefSeg3D: Enhancing Referring 3D Instance Segmentation via Structured Cross-Modal Graph Neural Networks.....	4551
<i>Zhipeng Qian, Yiwei Ma, Jiayi Ji, Xiaoshuai Sun</i>	
BARET: Balanced Attention Based Real Image Editing Driven by Target-Text Inversion.....	4560
<i>Yuming Qiao, Fanyi Wang, Jingwen Su, Yanhao Zhang, Yunjie Yu, Siyu Wu, Guo-Jun Qi</i>	
High-Fidelity 3D Head Avatars Reconstruction through Spatially-Varying Expression Conditioned Neural Radiance Field.....	4569
<i>Minghan Qin, Yifan Liu, Yuelang Xu, Xiaochen Zhao, Yebin Liu, Haoqian Wang</i>	
Text2City: One-Stage Text-Driven Urban Layout Regeneration.....	4578
<i>Yiming Qin, Nanxuan Zhao, Bin Sheng, Rynson W.H. Lau</i>	
Empowering CAM-Based Methods with Capability to Generate Fine-Grained and High-Faithfulness Explanations.....	4587
<i>Changqing Qiu, Fusheng Jin, Yining Zhang</i>	
High-Order Structure Based Middle-Feature Learning for Visible-Infrared Person Re-identification.....	4596
<i>Liuxiang Qiu, Si Chen, Yan Yan, Jing-Hao Xue, Da-Han Wang, Shunzhi Zhu</i>	
Mining Fine-Grained Image-Text Alignment for Zero-Shot Captioning via Text-Only Training.....	4605
<i>Longtian Qiu, Shan Ning, Xuming He</i>	
HiHPQ: Hierarchical Hyperbolic Product Quantization for Unsupervised Image Retrieval.....	4614
<i>Zexuan Qiu, Jiahong Liu, Yankai Chen, Irwin King</i>	
S2CycleDiff: Spatial-Spectral-Bilateral Cycle-Diffusion Framework for Hyperspectral Image Super-resolution.....	4623
<i>Jiahui Qu, Jie He, Wenqian Dong, Jingyu Zhao</i>	
E2HQV: High-Quality Video Generation from Event Camera via Theory-Inspired Model-Aided Deep Learning.....	4632
<i>Qiang Qu, Yiran Shen, Xiaoming Chen, Yuk Ying Chung, Tongliang Liu</i>	
BLiRF: Bandlimited Radiance Fields for Dynamic Scene Modeling.....	4641
<i>Sameera Ramasinghe, Violetta Shevchenko, Gil Avraham, Anton van den Hengel</i>	
Cross-Sentence Gloss Consistency for Continuous Sign Language Recognition.....	4650
<i>Qi Rao, Ke Sun, Xiaohan Wang, Qi Wang, Bang Zhang</i>	
Forecasting Bimanual Object Manipulation Sequences from Unimanual Observations.....	4659
<i>Haziq Razali, Yiannis Demiris</i>	
Multi-Step Denoising Scheduled Sampling: Towards Alleviating Exposure Bias for Diffusion Models.....	4667
<i>Zhiyao Ren, Yibing Zhan, Liang Ding, Gaoang Wang, Chaoyue Wang, Zhongyi Fan, Dacheng Tao</i>	
CRA-PCN: Point Cloud Completion with Intra- and Inter-level Cross-Resolution Transformers.....	4676
<i>Yi Rong, Haoran Zhou, Lixin Yuan, Cheng Mei, Jiahao Wang, Tong Lu</i>	
Entropic Open-Set Active Learning.....	4686
<i>Bardia Safaei, Vibashan VS, Celso M. de Melo, Vishal M. Patel</i>	

Generating Images of Rare Concepts Using Pre-trained Diffusion Models	4695
<i>Dvir Samuel, Rami Ben-Ari, Simon Raviv, Nir Darshan, Gal Chechik</i>	
RG-GAN: Dynamic Regenerative Pruning for Data-Efficient Generative Adversarial Networks	4704
<i>Divya Saxena, Jiannong Cao, Jiahao Xu, Tarun Kulshrestha</i>	
SeTformer Is What You Need for Vision and Language	4713
<i>Pourya Shamsolmoali, Masoumeh Zareapoor, Eric Granger, Michael Felsberg</i>	
Multi-Domain Multi-Scale Diffusion Model for Low-Light Image Enhancement	4722
<i>Kai Shang, Mingwen Shao, Chao Wang, Yuanshuo Cheng, Shuigen Wang</i>	
Polyper: Boundary Sensitive Polyp Segmentation	4731
<i>Hao Shao, Yang Zhang, Qibin Hou</i>	
Collaborative Consortium of Foundation Models for Open-World Few-Shot Learning	4740
<i>Shuai Shao, Yu Bai, Yan Wang, Baodi Liu, Bin Liu</i>	
FaceCoresetNet: Differentiable Coresets for Face Set Recognition	4748
<i>Gil Shapira, Yosi Keller</i>	
Decouple Content and Motion for Conditional Image-to-Video Generation.....	4757
<i>Cuifeng Shen, Yulu Gan, Chen Chen, Xiongwei Zhu, Lele Cheng, Tingting Gao, Jinzhi Wang</i>	
GroundVLP: Harnessing Zero-Shot Visual Grounding from Vision-Language Pre-training and Open-Vocabulary Object Detection.....	4766
<i>Haozhan Shen, Tiancheng Zhao, Mingwei Zhu, Jianwei Yin</i>	
Automatic Radiology Reports Generation via Memory Alignment Network	4776
<i>Hongyu Shen, Mingtao Pei, Juncai Liu, Zhaoxing Tian</i>	
CGMG: A Cross-Gaussian Mixture Generative Model for Few-Shot Semantic Segmentation.....	4784
<i>Junao Shen, Kun Kuang, Jiaheng Wang, Xinyu Wang, Tian Feng, Wei Zhang</i>	
Learn How to See: Collaborative Embodied Learning for Object Detection and Camera Adjusting.....	4793
<i>Lingdong Shen, Chunlei Huo, Nuo Xu, Chaowei Han, Zichen Wang</i>	
Distributed Manifold Hashing for Image Set Classification and Retrieval	4802
<i>Xiaobo Shen, Peizhuo Song, Yun-Hao Yuan, Yuhui Zheng</i>	
Controllable 3D Face Generation with Conditional Style Code Diffusion.....	4811
<i>Xiaolong Shen, Jianxin Ma, Chang Zhou, Zongxin Yang</i>	
Adaptive Integration of Partial Label Learning and Negative Learning for Enhanced Noisy Label Learning	4820
<i>Mengmeng Sheng, Zeren Sun, Zhenhuang Cai, Tao Chen, Yichao Zhou, Yazhou Yao</i>	
Transformer-Based No-Reference Image Quality Assessment via Supervised Contrastive Learning.....	4829
<i>Jinsong Shi, Pan Gao, Jie Qin</i>	
Explicit Visual Prompts for Visual Object Tracking.....	4838
<i>Liangtao Shi, Bineng Zhong, Qihua Liang, Ning Li, Shengping Zhang, Xianxian Li</i>	
Evidential Uncertainty-Guided Mitochondria Segmentation for 3D EM Images.....	4847
<i>Ruohua Shi, Lingyu Duan, Tiejun Huang, Tingting Jiang</i>	

Towards Squeezing-Averse Virtual Try-On via Sequential Deformation	4856
<i>Sang-Heon Shim, Jiwoo Chung, Jae-Pil Heo</i>	
DPA-P2PNet: Deformable Proposal-Aware P2PNet for Accurate Point-Based Cell Detection	4864
<i>Zhongyi Shui, Sunyi Zheng, Chenglu Zhu, Shichuan Zhang, Xiaoxuan Yu, Honglin Li, Jingxiong Li, Pingyi Chen, Lin Yang</i>	
DVANet: Disentangling View and Action Features for Multi-View Action Recognition.....	4873
<i>Nyle Siddiqui, Praveen Tirupattur, Mubarak Shah</i>	
Learning to Approximate Adaptive Kernel Convolution on Graphs	4882
<i>Jaeyoon Sim, Sooyeon Jeon, InJun Choi, Guorong Wu, Won Hwa Kim</i>	
Semi-supervised Active Learning for Video Action Detection.....	4891
<i>Ayush Singh, Aayush J Rana, Akash Kumar, Shruti Vyas, Yogesh Singh Rawat</i>	
DeblurSR: Event-Based Motion Deblurring under the Spiking Representation	4900
<i>Chen Song, Chandrajit Bajaj, Qixing Huang</i>	
Multi-Cross Sampling and Frequency-Division Reconstruction for Image Compressed Sensing	4909
<i>Heping Song, Jingyao Gong, Hongying Meng, Yuping Lai</i>	
Generalizable Fourier Augmentation for Unsupervised Video Object Segmentation	4918
<i>Huihui Song, Tiankang Su, Yuhui Zheng, Kaihua Zhang, Bo Liu, Dong Liu</i>	
Semantic-Aware Autoregressive Image Modeling for Visual Representation Learning.....	4925
<i>Kaiyou Song, Shan Zhang, Tong Wang</i>	
Self-Prompt Mechanism for Few-Shot Image Recognition.....	4934
<i>Mingchen Song, Huiqiang Wang, Guoqiang Zhong</i>	
Diverse Person: Customize Your Own Dataset for Text-Based Person Search	4943
<i>Zifan Song, Guosheng Hu, Cairong Zhao</i>	
V2Meow: Meowing to the Visual Beat via Video-to-Music Generation.....	4952
<i>Kun Su, Judith Yue Li, Qingqing Huang, Dima Kuzmin, Joonseok Lee, Chris Donahue, Fei Sha, Aren Jansen, Yu Wang, Mauro Verzetti, Timo Denk</i>	
F ³ -Pruning: A Training-Free and Generalized Pruning Strategy towards Faster and Finer Text-to-Video Synthesis	4961
<i>Sitong Su, Jianzhi Liu, Lianli Gao, Jingkuan Song</i>	
A Unified Environmental Network for Pedestrian Trajectory Prediction	4970
<i>Yuchao Su, Yuanman Li, Wei Wang, Jiantao Zhou, Xia Li</i>	
LRANet: Towards Accurate and Efficient Scene Text Detection with Low-Rank Approximation Network.....	4979
<i>Yuchen Su, Zhineng Chen, Zhiwen Shao, Yuning Du, Zhilong Ji, Jinfeng Bai, Yong Zhou, Yugang Jiang</i>	
Spatial-Semantic Collaborative Cropping for User Generated Content	4988
<i>Yukun Su, Yiwen Cao, Jingliang Deng, Fengyun Rao, Qingyao Wu</i>	
TR-DETR: Task-Reciprocal Transformer for Joint Moment Retrieval and Highlight Detection.....	4998
<i>Hao Sun, Mingyao Zhou, Wenjing Chen, Wei Xie</i>	

UniAP: Towards Universal Animal Perception in Vision via Few-Shot Learning	5008
<i>Meiqi Sun, Zhonghan Zhao, Wenhao Chai, Hanjun Luo, Shidong Cao, Yanting Zhang, Jenq-Neng Hwang, Gaoang Wang</i>	
CFR-ICL: Cascade-Forward Refinement with Iterative Click Loss for Interactive Image Segmentation	5017
<i>Shoukun Sun, Min Xian, Fei Xu, Luca Capriotti, Tiankai Yao</i>	
RL-SeqISP: Reinforcement Learning-Based Sequential Optimization for Image Signal Processing	5025
<i>Xinyu Sun, Zhikun Zhao, Lili Wei, Congyan Lang, Mingxuan Cai, Longfei Han, Juan Wang, Bing Li, Yuxuan Guo</i>	
PathAsst: A Generative Foundation AI Assistant towards Artificial General Intelligence of Pathology.....	5034
<i>Yuxuan Sun, Chenglu Zhu, Sunyi Zheng, Kai Zhang, Lin Sun, Zhongyi Shui, Yunlong Zhang, Honglin Li, Lin Yang</i>	
FG-EmoTalk: Talking Head Video Generation with Fine-Grained Controllable Facial Expressions.....	5043
<i>Zhaoxu Sun, Yuze Xuan, Fang Liu, Yang Xiang</i>	
Frequency-Aware Deepfake Detection: Improving Generalizability through Frequency Space Domain Learning.....	5052
<i>Chuangchuang Tan, Yao Zhao, Shikui Wei, Guanghua Gu, Ping Liu, Yunchao Wei</i>	
Compound Text-Guided Prompt Tuning via Image-Adaptive Cues.....	5061
<i>Hao Tan, Jun Li, Yizhuang Zhou, Jun Wan, Zhen Lei, Xiangyu Zhang</i>	
Occluded Person Re-identification via Saliency-Guided Patch Transfer	5070
<i>Lei Tan, Jiaer Xia, Wenfeng Liu, Pingyang Dai, Yongjian Wu, Liujuan Cao</i>	
Style2Talker: High-Resolution Talking Head Generation with Emotion Style and Art Style	5079
<i>Shuai Tan, Bin Ji, Ye Pan</i>	
Say Anything with Any Style	5088
<i>Shuai Tan, Bin Ji, Yu Ding, Ye Pan</i>	

Author Index