

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 6: AAAI Technical Tracks

- Computer Vision V

Vancouver, Canada
20 – 27 February 2024

Part 1 of 2

ISBN: 979-8-3313-0088-3

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 V

Semantic-Aware Data Augmentation for Text-To-Image Synthesis	5098
<i>Zhaorui Tan, Xi Yang, Kaizhu Huang</i>	
Data-Free Generalized Zero-Shot Learning	5108
<i>Bowen Tang, Jing Zhang, Long Yan, Qian Yu, Lu Sheng, Dong Xu</i>	
Offline and Online Optical Flow Enhancement for Deep Video Compression	5118
<i>Chuanbo Tang, Xihua Sheng, Zhuoyuan Li, Haotian Zhang, Li Li, Dong Liu</i>	
Manifold Constraints for Imperceptible Adversarial Attacks on Point Clouds.....	5127
<i>Keke Tang, Xu He, Weilong Peng, Jianpeng Wu, Yawen Shi, Daizong Liu, Pan Zhou, Wenping Wang, Zhihong Tian</i>	
Once and for All: Universal Transferable Adversarial Perturbation Against Deep Hashing-Based Facial Image Retrieval.....	5136
<i>Long Tang, Dengpan Ye, Yunna Lv, Chuanxi Chen, Yunming Zhang</i>	
Prior and Prediction Inverse Kernel Transformer for Single Image Defocus Deblurring	5145
<i>Peng Tang, Zhiqiang Xu, Chunlai Zhou, Pengfei Wei, Peng Han, Xin Cao, Tobias Lasser</i>	
Semantic Lens: Instance-Centric Semantic Alignment for Video Super-Resolution.....	5154
<i>Qi Tang, Yao Zhao, Meiqin Liu, Jian Jin, Chao Yao</i>	
Boosting Residual Networks with Group Knowledge.....	5162
<i>Shengji Tang, Peng Ye, Baopu Li, Weihao Lin, Tao Chen, Tong He, Chong Yu, Wanli Ouyang</i>	
Point-PEFT: Parameter-Efficient Fine-Tuning for 3D Pre-Trained Models	5171
<i>Yiwen Tang, Ray Zhang, Zoey Guo, Xianzheng Ma, Bin Zhao, Zhigang Wang, Dong Wang, Xuelong Li</i>	
Context-I2W: Mapping Images to Context-Dependent Words for Accurate Zero-Shot Composed Image Retrieval	5180
<i>Yuanmin Tang, Jing Yu, Keke Gai, Jiamin Zhuang, Gang Xiong, Yue Hu, Qi Wu</i>	
Generative-Based Fusion Mechanism for Multi-Modal Tracking.....	5189
<i>Zhangyong Tang, Tianyang Xu, Xiaojun Wu, Xue-Feng Zhu, Josef Kittler</i>	
Shadow Generation with Decomposed Mask Prediction and Attentive Shadow Filling.....	5198
<i>Xinhao Tao, Junyan Cao, Yan Hong, Li Niu</i>	
Towards Efficient and Effective Text-To-Video Retrieval with Coarse-To-Fine Visual Representation Learning.....	5207
<i>Kaibin Tian, Yanhua Cheng, Yi Liu, Xinglin Hou, Quan Chen, Han Li</i>	
Open-Vocabulary Video Relation Extraction.....	5215
<i>Wentao Tian, Zheng Wang, Yuqian Fu, Jingjing Chen, Lechao Cheng</i>	

Divide and Conquer: Hybrid Pre-Training for Person Search	5224
<i>Yanling Tian, Di Chen, Yunan Liu, Jian Yang, Shanshan Zhang</i>	
Taxonomy Driven Fast Adversarial Training.....	5233
<i>Kun Tong, Chengze Jiang, Jie Gui, Yuan Cao</i>	
End-To-End Real-Time Vanishing Point Detection with Transformer	5243
<i>Xin Tong, Shi Peng, Yufei Guo, Xuhui Huang</i>	
Discrete Cycle-Consistency Based Unsupervised Deep Graph Matching.....	5252
<i>Siddharth Tourani, Muhammad Haris Khan, Carsten Rother, Bogdan Savchynskyy</i>	
A Unified Masked Autoencoder with Patchified Skeletons for Motion Synthesis	5261
<i>Esteve Valls Mascaró, Hyemin Ahn, Dongheui Lee</i>	
CoVR: Learning Composed Video Retrieval from Web Video Captions	5270
<i>Lucas Ventura, Antoine Yang, Cordelia Schmid, Gül Varol</i>	
Supervision Interpolation Via LossMix: Generalizing Mixup for Object Detection and Beyond	5280
<i>Thanh Vu, Baochen Sun, Bodi Yuan, Alex Ngai, Yueqi Li, Jan-Michael Frahm</i>	
Integrated Decision Gradients: Compute Your Attributions Where the Model Makes Its Decision.....	5289
<i>Chase Walker, Sumit Jha, Kenny Chen, Rickard Ewetz</i>	
HISR: Hybrid Implicit Surface Representation for Photorealistic 3D Human Reconstruction	5298
<i>Angtian Wang, Yuanlu Xu, Nikolaos Sarafianos, Robert Maier, Edmond Boyer, Alan Yuille, Tony Tung</i>	
VIGC: Visual Instruction Generation and Correction	5309
<i>Bin Wang, Fan Wu, Xiao Han, Jiahui Peng, Huaping Zhong, Pan Zhang, Xiaoyi Dong, Weijia Li, Wei Li, Jiaqi Wang, Conghui He</i>	
Low-Light Face Super-Resolution Via Illumination, Structure, and Texture Associated Representation.....	5318
<i>Chenyang Wang, Junjun Jiang, Kui Jiang, Xianming Liu</i>	
SelfPromer: Self-Prompt Dehazing Transformers with Depth-Consistency	5327
<i>Cong Wang, Jinshan Pan, Wanyu Lin, Jiangxin Dong, Wei Wang, Xiao-Ming Wu</i>	
Correlation Matching Transformation Transformers for UHD Image Restoration.....	5336
<i>Cong Wang, Jinshan Pan, Wei Wang, Gang Fu, Siyuan Liang, Mengzhu Wang, Xiao-Ming Wu, Jun Liu</i>	
EulerMormer: Robust Eulerian Motion Magnification Via Dynamic Filtering Within Transformer	5345
<i>Fei Wang, Dan Guo, Kun Li, Meng Wang</i>	
Learning to Learn Better Visual Prompts	5354
<i>Fengxiang Wang, Wanrong Huang, Shaowu Yang, Qi Fan, Long Lan</i>	
MuST: Robust Image Watermarking for Multi-Source Tracing	5364
<i>Guanjie Wang, Zehua Ma, Chang Liu, Xi Yang, Han Fang, Weiming Zhang, Nenghai Yu</i>	
LION: Implicit Vision Prompt Tuning	5372
<i>Haixin Wang, Jianlong Chang, Yihang Zhai, Xiao Luo, Jinan Sun, Zhouchen Lin, Qi Tian</i>	

B-Spine: Learning B-Spline Curve Representation for Robust and Interpretable Spinal Curvature Estimation.....	5381
<i>Hao Wang, Qiang Song, Ruofeng Yin, Rui Ma</i>	
ViLT-CLIP: Video and Language Tuning CLIP with Multimodal Prompt Learning and Scenario-Guided Optimization	5390
<i>Hao Wang, Fang Liu, Licheng Jiao, Jiahao Wang, Zehua Hao, Shuo Li, Lingling Li, Puhua Chen, Xu Liu</i>	
Triple Feature Disentanglement for One-Stage Adaptive Object Detection.....	5401
<i>Haoan Wang, Shilong Jia, Tiejong Zeng, Guixu Zhang, Zhi Li</i>	
Neural Physical Simulation with Multi-Resolution Hash Grid Encoding	5410
<i>Haoxiang Wang, Tao Yu, Tianwei Yang, Hui Qiao, Qionghai Dai</i>	
Deep Unfolded Network with Intrinsic Supervision for Pan-Sharpener	5419
<i>Hebaixu Wang, Meiqi Gong, Xiaoguang Mei, Hao Zhang, Jiayi Ma</i>	
Continuous Piecewise-Affine Based Motion Model for Image Animation.....	5427
<i>Hexiang Wang, Fengqi Liu, Qianyu Zhou, Ran Yi, Xin Tan, Lizhuang Ma</i>	
Temporal Adaptive RGBT Tracking with Modality Prompt	5436
<i>Hongyu Wang, Xiaotao Liu, Yifan Li, Meng Sun, Dian Yuan, Jing Liu</i>	
SAUI: Scale-Aware Unseen Imagineer for Zero-Shot Object Detection.....	5445
<i>Jiahao Wang, Caixia Yan, Weizhan Zhang, Huan Liu, Hao Sun, Qinghua Zheng</i>	
Omnidirectional Image Super-Resolution Via Bi-Projection Fusion.....	5454
<i>Jiangang Wang, Yuning Cui, Yawen Li, Wenqi Ren, Xiaochun Cao</i>	
Adaptive FSS: A Novel Few-Shot Segmentation Framework Via Prototype Enhancement	5463
<i>Jing Wang, Jiangyun Li, Chen Chen, Yisi Zhang, Haoran Shen, Tianxiang Zhang</i>	
PointAttN: You Only Need Attention for Point Cloud Completion.....	5472
<i>Jun Wang, Ying Cui, Dongyan Guo, Junxia Li, Qingshan Liu, Chunhua Shen</i>	
EarthVQA: Towards Queryable Earth Via Relational Reasoning-Based Remote Sensing Visual Question Answering	5481
<i>Junjue Wang, Zhuo Zheng, Zihang Chen, Ailong Ma, Yanfei Zhong</i>	
Semi-Supervised Class-Agnostic Motion Prediction with Pseudo Label Regeneration and BEVMix.....	5490
<i>Kewei Wang, Yizheng Wu, Zhiyu Pan, Xingyi Li, Ke Xian, Zhe Wang, Zhiguo Cao, Guosheng Lin</i>	
Multi-Domain Incremental Learning for Face Presentation Attack Detection	5499
<i>Keyao Wang, Guosheng Zhang, Haixiao Yue, Ajian Liu, Gang Zhang, Haocheng Feng, Junyu Han, Errui Ding, Jingdong Wang</i>	
AltNeRF: Learning Robust Neural Radiance Field Via Alternating Depth-Pose Optimization	5508
<i>Kun Wang, Zhiqiang Yan, Huang Tian, Zhenyu Zhang, Xiang Li, Jun Li, Jian Yang</i>	
A Multimodal, Multi-Task Adapting Framework for Video Action Recognition	5517
<i>Mengmeng Wang, Jiazheng Xing, Boyuan Jiang, Jun Chen, Jianbiao Mei, Xingxing Zuo, Guang Dai, Jingdong Wang, Yong Liu</i>	
MsLPCC: A Multimodal-Driven Scalable Framework for Deep LiDAR Point Cloud Compression.....	5526
<i>Miaohui Wang, Runnan Huang, Hengjin Dong, Di Lin, Yun Song, Wuyuan Xie</i>	

Cycle-Consistency Learning for Captioning and Grounding	5535
<i>Ning Wang, Jiajun Deng, Mingbo Jia</i>	
Compositional Text-To-Image Synthesis with Attention Map Control of Diffusion Models	5544
<i>Ruichen Wang, Zekang Chen, Chen Chen, Jian Ma, Haonan Lu, Xiaodong Lin</i>	
AGS: Affordable and Generalizable Substitute Training for Transferable Adversarial Attack.....	5553
<i>Ruikui Wang, Yuanfang Guo, Yunhong Wang</i>	
DocNLC: A Document Image Enhancement Framework with Normalized and Latent Contrastive Representation for Multiple Degradations.....	5563
<i>Ruilu Wang, Yang Xue, Lianwen Jin</i>	
Towards Evidential and Class Separable Open Set Object Detection	5572
<i>Ruofan Wang, Rui-Wei Zhao, Xiaobo Zhang, Rui Feng</i>	
Suppressing Uncertainty in Gaze Estimation	5581
<i>Shijing Wang, Yaping Huang</i>	
What Effects the Generalization in Visual Reinforcement Learning: Policy Consistency with Truncated Return Prediction.....	5590
<i>Shuo Wang, Zhihao Wu, Xiaobo Hu, Jinwen Wang, Youfang Lin, Kai Lv</i>	
DeepAccident: A Motion and Accident Prediction Benchmark for V2X Autonomous Driving.....	5599
<i>Tianqi Wang, Sukmin Kim, Ji Wenxuan, Enze Xie, Chongjian Ge, Junsong Chen, Zhenguo Li, Ping Luo</i>	
Semantic-Guided Novel Category Discovery.....	5607
<i>Weishuai Wang, Ting Lei, Qingchao Chen, Yang Liu</i>	
HARDVS: Revisiting Human Activity Recognition with Dynamic Vision Sensors	5615
<i>Xiao Wang, Zongzhen Wu, Bo Jiang, Zhimin Bao, Lin Zhu, Guoqi Li, Yaowei Wang, Yonghong Tian</i>	
Structural Information Guided Multimodal Pre-Training for Vehicle-Centric Perception	5624
<i>Xiao Wang, Wentao Wu, Chenglong Li, Zhicheng Zhao, Zhe Chen, Yukai Shi, Jin Tang</i>	
ICAR: Image-Based Complementary Auto Reasoning	5633
<i>Xijun Wang, Anqi Liang, Junbang Liang, Ming Lin, Yu Lou, Shan Yang</i>	
GCNext: Towards the Unity of Graph Convolutions for Human Motion Prediction	5642
<i>Xinshun Wang, Qiongjie Cui, Chen Chen, Mengyuan Liu</i>	
CL2CM: Improving Cross-Lingual Cross-Modal Retrieval Via Cross-Lingual Knowledge Transfer	5651
<i>Yabing Wang, Fan Wang, Jianfeng Dong, Hao Luo</i>	
OSFFNet: Omni-Stage Feature Fusion Network for Lightweight Image Super-Resolution	5660
<i>Yang Wang, Tao Zhang</i>	
Prompting Segmentation with Sound is Generalizable Audio-Visual Source Localizer	5669
<i>Yaoting Wang, Weisong Liu, Guangyao Li, Jian Ding, Di Hu, Xi Li</i>	
Mask-Homo: Pseudo Plane Mask-Guided Unsupervised Multi-Homography Estimation	5678
<i>Yasi Wang, Hong Liu, Chao Zhang, Lu Xu, Qiang Wang</i>	
PointPatchMix: Point Cloud Mixing with Patch Scoring	5686
<i>Yi Wang, Jiaze Wang, Jinpeng Li, Zixu Zhao, Guangyong Chen, Anfeng Liu, Pheng Ann Heng</i>	

Data Distribution Distilled Generative Model for Generalized Zero-Shot Recognition	5695
<i>Yijie Wang, Mingjian Hong, Luwen Huangfu, Sheng Huang</i>	
SiMA-Hand: Boosting 3D Hand-Mesh Reconstruction by Single-To-Multi-View Adaptation	5704
<i>Yinqiao Wang, Hao Xu, Pheng Ann Heng, Chi-Wing Fu</i>	
SQLdepth: Generalizable Self-Supervised Fine-Structured Monocular Depth Estimation.....	5713
<i>Youhong Wang, Yunji Liang, Hao Xu, Shaohui Jiao, Hongkai Yu</i>	
H2GFormer: Horizontal-To-Global Voxel Transformer for 3D Semantic Scene Completion.....	5722
<i>Yu Wang, Chao Tong</i>	
Exploring Diverse Representations for Open Set Recognition.....	5731
<i>Yu Wang, Junxian Mu, Pengfei Zhu, Qinghua Hu</i>	
SMILEtrack: SiMilarity LEarning for Occlusion-Aware Multiple Object Tracking.....	5740
<i>Yu-Hsiang Wang, Jun-Wei Hsieh, Ping-Yang Chen, Ming-Ching Chang, Hung-Hin So, Xin Li</i>	
Learning Hierarchical Prompt with Structured Linguistic Knowledge for Vision-Language Models	5749
<i>Yubin Wang, Xinyang Jiang, De Cheng, Dongsheng Li, Cairong Zhao</i>	
TOP-ReID: Multi-Spectral Object Re-Identification with Token Permutation	5758
<i>Yuhao Wang, Xuehu Liu, Pingping Zhang, Hu Lu, Zhengzheng Tu, Huchuan Lu</i>	
GMMFormer: Gaussian-Mixture-Model Based Transformer for Efficient Partially Relevant Video Retrieval	5767
<i>Yuting Wang, Jinpeng Wang, Bin Chen, Ziyun Zeng, Shu-Tao Xia</i>	
Out of Thin Air: Exploring Data-Free Adversarial Robustness Distillation	5776
<i>Yuzheng Wang, Zhaoyu Chen, Dingkang Yang, Pinxue Guo, Kaixun Jiang, Wenqiang Zhang, Lizhe Qi</i>	
QAGait: Revisit Gait Recognition from a Quality Perspective	5785
<i>Zengbin Wang, Saihui Hou, Man Zhang, Xu Liu, Chunshui Cao, Yongzhen Huang, Peipei Li, Shibiao Xu</i>	
Enhancing Hyperspectral Images Via Diffusion Model and Group-Autoencoder Super-Resolution Network.....	5794
<i>Zhaoyang Wang, Dongyang Li, Mingyang Zhang, Hao Luo, Maoguo Gong</i>	

PART 2

SkyScript: A Large and Semantically Diverse Vision-Language Dataset for Remote Sensing	5805
<i>Zhecheng Wang, Rajanie Prabha, Tianyuan Huang, Jiajun Wu, Ram Rajagopal</i>	
DTMFormer: Dynamic Token Merging for Boosting Transformer-Based Medical Image Segmentation	5814
<i>Zehao Wang, Xian Lin, Nannan Wu, Li Yu, Kwang-Ting Cheng, Zengqiang Yan</i>	
SGNet: Structure Guided Network Via Gradient-Frequency Awareness for Depth Map Super- Resolution.....	5823
<i>Zhengxue Wang, Zhiqiang Yan, Jian Yang</i>	
Vision Transformer Off-The-Shelf: A Surprising Baseline for Few-Shot Class-Agnostic Counting	5832
<i>Zhicheng Wang, Liwen Xiao, Zhiguo Cao, Hao Lu</i>	

Existence is Chaos: Enhancing 3D Human Motion Prediction with Uncertainty Consideration	5841
<i>Zhihao Wang, Yulin Zhou, Ningyu Zhang, Xiaosong Yang, Jun Xiao, Zhao Wang</i>	
Heterogeneous Test-Time Training for Multi-Modal Person Re-Identification.....	5850
<i>Zi Wang, Huaibo Huang, Aihua Zheng, Ran He</i>	
Fine-Grained Prototypes Distillation for Few-Shot Object Detection.....	5859
<i>Zichen Wang, Bo Yang, Haonan Yue, Zhenghao Ma</i>	
Semantic Complete Scene Forecasting from a 4D Dynamic Point Cloud Sequence.....	5867
<i>Zifan Wang, Zhuorui Ye, Haoran Wu, Junyu Chen, Li Yi</i>	
Enhanced Fine-Grained Motion Diffusion for Text-Driven Human Motion Synthesis.....	5876
<i>Dong Wei, Xiaoning Sun, Huaijiang Sun, Shengxiang Hu, Bin Li, Weiqing Li, Jianfeng Lu</i>	
Image as a Language: Revisiting Scene Text Recognition Via Balanced, Unified and Synchronized Vision-Language Reasoning Network.....	5885
<i>Jiajun Wei, Hongjian Zhan, Yue Lu, Xiao Tu, Bing Yin, Cong Liu, Umapada Pal</i>	
WeakPCSOD: Overcoming the Bias of Box Annotations for Weakly Supervised Point Cloud Salient Object Detection.....	5894
<i>Jun Wei, S. Kevin Zhou, Shuguang Cui, Zhen Li</i>	
RetouchFormer: Semi-Supervised High-Quality Face Retouching Transformer with Prior-Based Selective Self-Attention.....	5903
<i>Xue Wen, Lianxin Xie, Le Jiang, Tianyi Chen, Si Wu, Cheng Liu, Hau-San Wong</i>	
Mean Teacher DETR with Masked Feature Alignment: A Robust Domain Adaptive Detection Transformer Framework.....	5912
<i>Weixi Weng, Chun Yuan</i>	
Keep the Faith: Faithful Explanations in Convolutional Neural Networks for Case-Based Reasoning.....	5921
<i>Tom Nuno Wolf, Fabian Bongratz, Anne-Marie Rickmann, Sebastian Pölsterl, Christian Wachinger</i>	
Factorized Diffusion Autoencoder for Unsupervised Disentangled Representation Learning	5930
<i>Ancong Wu, Wei-Shi Zheng</i>	
3D-STMN: Dependency-Driven Superpoint-Text Matching Network for End-To-End 3D Referring Expression Segmentation	5940
<i>Changli Wu, Yiwei Ma, Qi Chen, Haowei Wang, Gen Luo, Jiayi Ji, Xiaoshuai Sun</i>	
SCD-Net: Spatiotemporal Clues Disentanglement Network for Self-Supervised Skeleton-Based Action Recognition.....	5949
<i>Cong Wu, Xiao-Jun Wu, Josef Kittler, Tianyang Xu, Sara Ahmed, Muhammad Awais, Zhenhua Feng</i>	
G-NAS: Generalizable Neural Architecture Search for Single Domain Generalization Object Detection	5958
<i>Fan Wu, Jinling Gao, Lanqing Hong, Xinbing Wang, Chenghu Zhou, Nanyang Ye</i>	
Multiscale Low-Frequency Memory Network for Improved Feature Extraction in Convolutional Neural Networks.....	5967
<i>Fuzhi Wu, Jiasong Wu, Youyong Kong, Chunfeng Yang, Guanyu Yang, Huazhong Shu, Guy Carrault, Lotfi Senhadji</i>	

Learning from History: Task-Agnostic Model Contrastive Learning for Image Restoration	5976
<i>Gang Wu, Junjun Jiang, Kui Jiang, Xianming Liu</i>	
Hybrid-Supervised Dual-Search: Leveraging Automatic Learning for Loss-Free Multi-Exposure Image Fusion	5985
<i>Guanyao Wu, Hongming Fu, Jinyuan Liu, Long Ma, Xin Fan, Risheng Liu</i>	
When to Grow? a Fitting Risk-Aware Policy for Layer Growing in Deep Neural Networks.....	5994
<i>Haihong Wu, Wei Wang, Tamasha Malepathirana, Damith Senanayake, Denny Oetomo, Saman Halgamuge</i>	
P-Laplacian Adaptation for Generative Pre-Trained Vision-Language Models	6003
<i>Haoyuan Wu, Xinyun Zhang, Peng Xu, Peiyu Liao, Xufeng Yao, Bei Yu</i>	
Task-Adaptive Prompted Transformer for Cross-Domain Few-Shot Learning	6012
<i>Jiamin Wu, Xin Liu, Xiaotian Yin, Tianzhu Zhang, Yongdong Zhang</i>	
SyFormer: Structure-Guided Synergism Transformer for Large-Portion Image Inpainting.....	6021
<i>Jie Wu, Yuchao Feng, Honghui Xu, Chuanmeng Zhu, Jianwei Zheng</i>	
MedSegDiff-V2: Diffusion-Based Medical Image Segmentation with Transformer	6030
<i>Junde Wu, Wei Ji, Huazhu Fu, Min Xu, Yueming Jin, Yanwu Xu</i>	
Selective and Orthogonal Feature Activation for Pedestrian Attribute Recognition	6039
<i>Junyi Wu, Yan Huang, Min Gao, Yuzhen Niu, Mingjing Yang, Zhipeng Gao, Jianqiang Zhao</i>	
Swift-Mapping: Online Neural Implicit Dense Mapping in Urban Scenes	6048
<i>Ke Wu, Kaizhao Zhang, Mingzhe Gao, Jieru Zhao, Zhongxue Gan, Wenchao Ding</i>	
CPN: Complementary Proposal Network for Unconstrained Text Detection.....	6057
<i>Longhuang Wu, Shangxuan Tian, Youxin Wang, Pengfei Xiong</i>	
Toward Open-Set Human Object Interaction Detection.....	6066
<i>Mingrui Wu, Yuqi Liu, Jiayi Ji, Xiaoshuai Sun, Rongrong Ji</i>	
VadCLIP: Adapting Vision-Language Models for Weakly Supervised Video Anomaly Detection.....	6074
<i>Peng Wu, Xuerong Zhou, Guansong Pang, Lingru Zhou, Qingsen Yan, Peng Wang, Yanning Zhang</i>	
Temporal Correlation Vision Transformer for Video Person Re-Identification.....	6083
<i>Pengfei Wu, Le Wang, Sanping Zhou, Gang Hua, Changyin Sun</i>	
Point-To-Spike Residual Learning for Energy-Efficient 3D Point Cloud Classification.....	6092
<i>Qiaoyun Wu, Quanxiao Zhang, Chunyu Tan, Yun Zhou, Changyin Sun</i>	
Segment Beyond View: Handling Partially Missing Modality for Audio-Visual Semantic Segmentation	6100
<i>Renjie Wu, Hu Wang, Feras Dayoub, Hsiang-Ting Chen</i>	
Towards Transferable Adversarial Attacks with Centralized Perturbation	6109
<i>Shangbo Wu, Yu-an Tan, Yajie Wang, Ruinan Ma, Wencong Ma, Yuanzhang Li</i>	
CLIM: Contrastive Language-Image Mosaic for Region Representation	6117
<i>Size Wu, Wenwei Zhang, Lumin Xu, Sheng Jin, Wentao Liu, Chen Change Loy</i>	
SphereDiffusion: Spherical Geometry-Aware Distortion Resilient Diffusion Model.....	6126
<i>Tao Wu, Xuwei Li, Zhongang Qi, Di Hu, Xintao Wang, Ying Shan, Xi Li</i>	

LRS: Enhancing Adversarial Transferability Through Lipschitz Regularized Surrogate	6135
<i>Tao Wu, Tie Luo, Donald C. Wunsch II</i>	
CR-SAM: Curvature Regularized Sharpness-Aware Minimization	6144
<i>Tao Wu, Tie Luo, Donald C. Wunsch II</i>	
Semi-Supervised 3D Object Detection with PatchTeacher and PillarMix.....	6153
<i>Xiaopei Wu, Liang Peng, Liang Xie, Yuenan Hou, Binbin Lin, Xiaoshui Huang, Haifeng Liu, Deng Cai, Wanli Ouyang</i>	
Text-Based Occluded Person Re-Identification Via Multi-Granularity Contrastive Consistency Learning	6162
<i>Xinyi Wu, Wentao Ma, Dan Guo, Tongqing Zhou, Shan Zhao, Zhiping Cai</i>	
CMG-Net: Robust Normal Estimation for Point Clouds Via Chamfer Normal Distance and Multi-Scale Geometry	6171
<i>Yingrui Wu, Mingyang Zhao, Keqiang Li, Weize Quan, Tianqi Yu, Jianfeng Yang, Xiaohong Jia, Dong-Ming Yan</i>	
WaveFormer: Wavelet Transformer for Noise-Robust Video Inpainting.....	6180
<i>Zhiliang Wu, Changchang Sun, Hanyu Xuan, Gaowen Liu, Yan Yan</i>	
FD3D: Exploiting Foreground Depth Map for Feature-Supervised Monocular 3D Object Detection.....	6189
<i>Zizhang Wu, Yuanzhu Gan, Yunzhe Wu, Ruihao Wang, Xiaoquan Wang, Jian Pu</i>	
Attention Disturbance and Dual-Path Constraint Network for Occluded Person Re-Identification.....	6198
<i>Jiaer Xia, Lei Tan, Pingyang Dai, Mingbo Zhao, Yongjian Wu, Liujuan Cao</i>	
Locality Preserving Refinement for Shape Matching with Functional Maps.....	6207
<i>Yifan Xia, Yifan Lu, Yuan Gao, Jiayi Ma</i>	
SocialCVAE: Predicting Pedestrian Trajectory Via Interaction Conditioned Latents	6216
<i>Wei Xiang, Haoteng YIN, He Wang, Xiaogang Jin</i>	
Dynamic Semantic-Based Spatial Graph Convolution Network for Skeleton-Based Human Action Recognition	6225
<i>Jianyang Xie, Yanda Meng, Yitian Zhao, Anh Nguyen, Xiaoyun Yang, Yalin Zheng</i>	
G2P-DDM: Generating Sign Pose Sequence from Gloss Sequence with Discrete Diffusion Model.....	6234
<i>Pan Xie, Qipeng Zhang, Peng Taiying, Hao Tang, Yao Du, Zexian Li</i>	
Towards Understanding Future: Consistency Guided Probabilistic Modeling for Action Anticipation	6243
<i>Zhao Xie, Yadong Shi, Kewei Wu, Yaru Cheng, Dan Guo</i>	
Towards Detailed Text-To-Motion Synthesis Via Basic-To-Advanced Hierarchical Diffusion Model	6252
<i>Zhenyu Xie, Yang Wu, Xuehao Gao, Zhongqian Sun, Wei Yang, Xiaodan Liang</i>	
Learning by Erasing: Conditional Entropy Based Transferable Out-Of-Distribution Detection.....	6261
<i>Meng Xing, Zhiyong Feng, Yong Su, Changjae Oh</i>	
Unsupervised Action Segmentation Via Fast Learning of Semantically Consistent Actoms	6270
<i>Zheng Xing, Weibing Zhao</i>	
SPEAL: Skeletal Prior Embedded Attention Learning for Cross-Source Point Cloud Registration	6279
<i>Kezheng Xiong, Maoji Zheng, Qingshan Xu, Chenglu Wen, Siqi Shen, Cheng Wang</i>	

Patched Line Segment Learning for Vector Road Mapping	6288
<i>Jiakun Xu, Bowen Xu, Gui-Song Xia, Liang Dong, Nan Xue</i>	
MuLTI: Efficient Video-And-Language Understanding with Text-Guided MultiWay-Sampler and Multiple Choice Modeling	6297
<i>Jiaqi Xu, Bo Liu, Yunkuo Chen, Mengli Cheng, Xing Shi</i>	
Regulating Intermediate 3D Features for Vision-Centric Autonomous Driving.....	6306
<i>Junkai Xu, Liang Peng, Haoran Cheng, Linxuan Xia, Qi Zhou, Dan Deng, Wei Qian, Wenxiao Wang, Deng Cai</i>	
ZOOM: Learning Video Mirror Detection with Extremely-Weak Supervision.....	6315
<i>Ke Xu, Tsun Wai Siu, Rynson W.H. Lau</i>	
Weakly Supervised Multimodal Affordance Grounding for Egocentric Images	6324
<i>Lingjing Xu, Yang Gao, Wenfeng Song, Aimin Hao</i>	
Gaze from Origin: Learning for Generalized Gaze Estimation by Embedding the Gaze Frontalization Process	6333
<i>Mingjie Xu, Feng Lu</i>	
HACDR-Net: Heterogeneous-Aware Convolutional Network for Diabetic Retinopathy Multi-Lesion Segmentation	6342
<i>QiHao Xu, Xiaoling Luo, Chao Huang, Chengliang Liu, Jie Wen, Jialei Wang, Yong Xu</i>	
Learning Invariant Inter-Pixel Correlations for Superpixel Generation	6351
<i>Sen Xu, Shikui Wei, Tao Ruan, Lixin Liao</i>	
Direction-Aware Video Demoiréing with Temporal-Guided Bilateral Learning.....	6360
<i>Shuning Xu, Binbin Song, Xiangyu Chen, Jiantao Zhou</i>	
Spectral Prompt Tuning: Unveiling Unseen Classes for Zero-Shot Semantic Segmentation.....	6369
<i>Wenhao Xu, Rongtao Xu, Changwei Wang, Shibiao Xu, Li Guo, Man Zhang, Xiaopeng Zhang</i>	
SCTNet: Single-Branch CNN with Transformer Semantic Information for Real-Time Segmentation.....	6378
<i>Zhengze Xu, Dongyue Wu, Changqian Yu, Xiangxiang Chu, Nong Sang, Changxin Gao</i>	
Chain of Generation: Multi-Modal Gesture Synthesis Via Cascaded Conditional Control.....	6387
<i>Zunnan Xu, Yachao Zhang, Sicheng Yang, Ronghui Li, Xiu Li</i>	
Decoupled Contrastive Learning for Long-Tailed Recognition	6396
<i>Shiyu Xuan, Shiliang Zhang</i>	
Revisiting Gradient Pruning: A Dual Realization for Defending Against Gradient Attacks.....	6404
<i>Lulu Xue, Shengshan Hu, Ruizhi Zhao, Leo Yu Zhang, Shengqing Hu, Lichao Sun, Dezhong Yao</i>	
A Convolutional Neural Network Interpretable Framework for Human Ventral Visual Pathway Representation	6413
<i>Mufan Xue, Xinyu Wu, Jinlong Li, Xuesong Li, Guoyuan Yang</i>	
Self-Supervised 3D Human Mesh Recovery from a Single Image with Uncertainty-Aware Learning.....	6422
<i>Guoli Yan, Zichun Zhong, Jing Hua</i>	
HORIZON: High-Resolution Semantically Controlled Panorama Synthesis.....	6431
<i>Kun Yan, Lei Ji, Chenfei Wu, Jian Liang, Ming Zhou, Nan Duan, Shuai Ma</i>	

CF-NeRF: Camera Parameter Free Neural Radiance Fields with Incremental Learning	6440
<i>Qingsong Yan, Qiang Wang, Kaiyong Zhao, Jie Chen, Bo Li, Xiaowen Chu, Fei Deng</i>	
Referred by Multi-Modality: A Unified Temporal Transformer for Video Object Segmentation.....	6449
<i>Shilin Yan, Renrui Zhang, Ziyu Guo, Wenchao Chen, Wei Zhang, Hongyang Li, Yu Qiao, Hao Dong, Zhongjiang He, Peng Gao</i>	
Embracing Language Inclusivity and Diversity in CLIP Through Continual Language Learning.....	6458
<i>Bang Yang, Yong Dai, Xuxin Cheng, Yaowei Li, Asif Raza, Yuxian Zou</i>	
Geometry-Guided Domain Generalization for Monocular 3D Object Detection	6467
<i>Fan Yang, Hui Chen, Yuwei He, Sicheng Zhao, Chenghao Zhang, Kai Ni, Guiguang Ding</i>	
Diversity-Authenticity Co-Constrained Stylization for Federated Domain Generalization in Person Re-Identification.....	6477
<i>Fengxiang Yang, Zhun Zhong, Zhiming Luo, Yifan He, Shaozi Li, Nicu Sebe</i>	
Semantic-Aware Transformation-Invariant RoI Align.....	6486
<i>Guo-Ye Yang, George Kiyohiro Nakayama, Zi-Kai Xiao, Tai-Jiang Mu, Xiaolei Huang, Shi-Min Hu</i>	
FACL-Attack: Frequency-Aware Contrastive Learning for Transferable Adversarial Attacks	6494
<i>Hunmin Yang, Jongoh Jeong, Kuk-Jin Yoon</i>	

Author Index