



We've just launched a new service: our brand new **dblp SPARQL query service**. Read more about it in our latest blog post or try out some of the SPARQL queries linked on the dblp web pages below.

X



[+] CVPR 2023: Vancouver, BC, Canada

> Home > Conferences and Workshops > CVPR

Dagstuhl

IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2023, Vancouver, BC, Canada, June 17-24, 2023. IEEE 2023, ISBN 979-8-3503-0129-8

Shikhar Bahl, Russell Mendonca, Lili Chen, Unnat Jain, Deepak Pathak: **Affordances from Human Videos as a Versatile Representation for Robotics.** 1-13

Lei Jin, Gen Luo, Yiyi Zhou, Xiaoshuai Sun, Guannan Jiang, Annan Shu, Rongrong Ji: **RefCLIP: A Universal Teacher for Weakly Supervised Referring Expression Comprehension.** 1-10

Adithya Pediredla, Srinivasa G. Narasimhan , Maysamreza Chamanzar, Ioannis Gkioulekas : **Megahertz Light Steering Without Moving Parts.** 1-12

Yu-Lun Liu, Chen Gao, Andreas Meuleman, Hung-Yu Tseng, Ayush Saraf, Changil Kim, Yung-Yu Chuang, Johannes Kopf, Jia-Bin Huang: **Robust Dynamic Radiance Fields.** 13-23

Yu Chen, Gim Hee Lee: **DBARF: Deep Bundle-Adjusting Generalizable Neural Radiance Fields.** 24-34

Bingfan Zhu, Yanchao Yang, Xulong Wang , Youyi Zheng, Leonidas J. Guibas: **VDN-NeRF: Resolving Shape-Radiance Ambiguity via View-Dependence Normalization.** 35-45

Yifan Jiang, Peter Hedman, Ben Mildenhall, Dejia Xu, Jonathan T. Barron, Zhangyang Wang, Tianfan Xue: **AligNeRF: High-Fidelity Neural Radiance Fields via Alignment-Aware Training.** 46-55

Deborah Levy , Amit Peleg, Naama Pearl, Dan Rosenbaum, Derya Akkaynak , Simon Korman, Tali Treibitz: **SeaThru-NeRF: Neural Radiance Fields in Scattering Media.** 56-65

Brian K. S. Isaac-Medina, Chris G. Willcocks , Toby P. Breckon: **Exact-NeRF: An Exploration of a Precise Volumetric Parameterization for Neural Radiance Fields.** 66-75

Liao Wang, Qiang Hu, Qihan He, Ziyu Wang, Jingyi Yu, Tinne Tuytelaars , Lan Xu, Minye Wu: **Neural Residual Radiance Fields for Streamably Free-Viewpoint Videos.** 76-87

Han Yan, Celong Liu, Chao Ma, Xing Mei: **Plen-VDB: Memory Efficient VDB-Based Radiance Fields for Fast Training and Rendering.** 88-96

Xin Huang, Qi Zhang, Ying Feng, Xiaoyu Li, Xuan Wang, Qing Wang: **Local Implicit Ray Function for Generalizable Radiance Field Representation.** 97-107

Yiming Gao, Yan-Pei Cao, Ying Shan: **SurfelNeRF: Neural Surfel Radiance Fields for Online Photorealistic Reconstruction of Indoor Scenes.** 108-118

Yi Zhang, Xiaoyang Huang, Bingbing Ni, Wenjun Zhang, Teng Li: **Frequency-Modulated Point Cloud Rendering with Easy Editing.** 119-129

Ang Cao, Justin Johnson: **HexPlane: A Fast Representation for Dynamic Scenes.** 130-141

Markus Worchel, Marc Alexa: **Differentiable Shadow Mapping for Efficient Inverse Graphics.** 142-153

Peng Dai, Yinda Zhang, Xin Yu, Xiaoyang Lyu, Xiaojuan Qi: **Hybrid Neural Rendering for Large-Scale Scenes with Motion Blur.** 154-164

Haian Jin, Isabella Liu, Peijia Xu, Xiaoshuai Zhang, Songfang Han, Sai Bi, Xiaowei Zhou, Zexiang Xu, Hao Su: **TensoIR: Tensorial Inverse Rendering.** 165-174

[+] SPARQL queries

[+] Refine list

showing all 2,357 records

refine by search term

refine by author

Yu Qiao (24)
Ying Shan (18)
Luc Van Gool (15)
Zhiwei Xiong (14)
Leonidas J. Guibas (14)
Yi-Zhe Song (14)
Hongsheng Li (14)
Dacheng Tao (14)
Yi Yang (13)
Michael J. Black (13)

8,810 more options

refine by access

closed (2,353)
withdrawn (3)



- Jingwang Ling, Zhibo Wang, Feng Xu:
ShadowNeuS: Neural SDF Reconstruction by Shadow Ray Supervision. 175-185
- S. Mahdi H. Miangoleh, Zoya Bylinskii, Eric Kee, Eli Shechtman, Yagiz Aksoy :
Realistic Saliency Guided Image Enhancement. 186-194
- Yiqun Mei, He Zhang, Xuaner Zhang, Jianming Zhang, Zhixin Shu, Yilin Wang, Zijun Wei, Shi Yan, Hyunjoon Jung, Vishal M. Patel:
LightPainter: Interactive Portrait Relighting with Freehand Scribble. 195-205
- Xianmin Xu, Yuxin Lin, Haoyang Zhou, Chong Zeng , Yixin Yu, Kun Zhou, Hongzhi Wu:
A Unified Spatial-Angular Structured Light for Single-View Acquisition of Shape and Reflectance. 206-215
- Ruichen Zheng, Peng Li, Haoqian Wang, Tao Yu:
Learning Visibility Field for Detailed 3D Human Reconstruction and Relighting. 216-226
- Junborg Jang, Kwonmoo Lee , Tae-Kyun Kim :
Unsupervised Contour Tracking of Live Cells by Mechanical and Cycle Consistency Losses. 227-236
- Yu-Tao Liu, Li Wang, Jie Yang, Weikai Chen, Xiaoxu Meng, Bo Yang, Lin Gao:
NeUDF: Leaning Neural Unsigned Distance Fields with Volume Rendering. 237-247
- Xiaoxu Meng, Weikai Chen, Bo Yang:
NeAT: Learning Neural Implicit Surfaces with Arbitrary Topologies from Multi-View Images. 248-258
- Zhen Wang, Shijie Zhou, Jeong Joon Park, Despoina Paschalidou , Suya You, Gordon Wetzstein, Leonidas J. Guibas, Achuta Kadambi:
ALTO: Alternating Latent Topologies for Implicit 3D Reconstruction. 259-270
- Zhaoyang Lyu, Jinyi Wang, Yuwei An, Ya Zhang , Dahua Lin, Bo Dai:
Controllable Mesh Generation Through Sparse Latent Point Diffusion Models. 271-280
- Ke Li, Kaiyue Pang, Yi-Zhe Song:
Photo Pre-Training, But for Sketch. 275-285
- Simon Weber, Nikolaus Demmel, Tin Chon Chan, Daniel Cremers:
Power Bundle Adjustment for Large-Scale 3D Reconstruction. 281-289
- Aayush Bansal, Michael Zollhöfer:
Neural Pixel Composition for 3D-4D View Synthesis from Multi-Views. 290-299
- Chen-Hsuan Lin, Jun Gao, Luming Tang, Towaki Takikawa, Xiaohui Zeng, Xun Huang, Karsten Kreis, Sanja Fidler, Ming-Yu Liu, Tsung-Yi Lin:
Magic3D: High-Resolution Text-to-3D Content Creation. 300-309
- Li Ma, Xiaoyu Li, Jing Liao , Pedro V. Sander:
3D Video Loops from Asynchronous Input. 310-320
- Jiaxin Xie, Hao Ouyang, Jingtian Piao, Chenyang Lei, Qifeng Chen:
High-fidelity 3D GAN Inversion by Pseudo-multi-view Optimization. 321-331
- Leheng Li, Qing Lian, Luozhou Wang, Ningning Ma, Yingcong Chen:
Lift3D: Synthesize 3D Training Data by Lifting 2D GAN to 3D Generative Radiance Field. 332-341
- Fei Yin, Yong Zhang, Xuan Wang, Tengfei Wang, Xiaoyu Li, Yuan Gong, Yanbo Fan, Xiaodong Cun, Ying Shan, Cengiz Öztireli, Yujiu Yang :
3D GAN Inversion with Facial Symmetry Prior. 342-351
- Diqiong Jiang, Dan Song, Ruofeng Tong, Min Tang:
StyleIPSB: Identity-Preserving Semantic Basis of StyleGAN for High Fidelity Face Swapping. 352-361
- Haoran Bai, Di Kang, Haoxian Zhang, Jinshan Pan, Linchao Bao :
FFHQ-UV: Normalized Facial UV-Texture Dataset for 3D Face Reconstruction. 362-371
- Chunlu Li , Andreas Morel-Forster, Thomas Vetter, Bernhard Egger , Adam Kortylewski:
Robust Model-based Face Reconstruction through Weakly-Supervised Outlier Segmentation. 372-381
- Zhenyu Zhang, Renwang Chen, Weijian Cao, Ying Tai, Chengjie Wang:
Learning Neural Proto-Face Field for Disentangled 3D Face Modeling in the Wild. 382-393
- Biwen Lei, Jianqiang Ren, Mengyang Feng, Miaomiao Cui, Xuansong Xie:
A Hierarchical Representation Network for Accurate and Detailed Face Reconstruction from In-The-Wild Images. 394-403
- Kacper Kania, Stephan J. Garbin, Andrea Tagliasacchi, Virginia Estellers, Kwang Moo Yi, Julien Valentin, Tomasz Trzcinski, Marek Kowalski:
BlendFields: Few-Shot Example-Driven Facial Modeling. 404-415
- Chuhan Chen, Matthew O'Toole, Gaurav Bharaj, Pablo Garrido:
Implicit Neural Head Synthesis via Controllable Local Deformation Fields. 416-426
- Youxin Pang, Yong Zhang, Weize Quan, Yanbo Fan, Xiaodong Cun, Ying Shan, Dong-Ming Yan:
DPE: Disentanglement of Pose and Expression for General Video Portrait Editing. 427-436
- Sijing Wu , Yichao Yan, Yunhao Li, Yuhao Cheng, Wenhan Zhu, Ke Gao, Xiaobo Li, Guangtao Zhai:
GANHead: Towards Generative Animatable Neural Head Avatars. 437-447



- Jonathan Tseng, Rodrigo Castellon, C. Karen Liu :
EDGE: Editable Dance Generation From Music. 448-458
- Aliaksandr Siarohin, Willi Menapace, Ivan Skorokhodov, Kyle Olszewski, Jian Ren, Hsin-Ying Lee, Menglei Chai, Sergey Tulyakov:
Unsupervised Volumetric Animation. 458-469
- Hugo Bertiche, Niloy J. Mitra, Kuldeep Kulkarni, Chun-Hao Paul Huang, Tuanfeng Y. Wang, Meysam Madadi, Sergio Escalera , Duygu Ceylan:
Blowing in the Wind: CycleNet for Human Cinemagraphs from Still Images. 459-468
- Hongwei Yi, Hualin Liang, Yifei Liu, Qiong Cao, Yandong Wen, Timo Bolkart, Dacheng Tao, Michael J. Black:
Generating Holistic 3D Human Motion from Speech. 469-480
- Yuming Du, Robin Kips, Albert Pumarola, Sebastian Starke, Ali K. Thabet, Artsiom Sanakoyeu:
Avatars Grow Legs: Generating Smooth Human Motion from Sparse Tracking Inputs with Diffusion Model. 481-490
- Fang Zhao, Zekun Li, Shaoli Huang, Junwu Weng, Tianfei Zhou, Guo-Sen Xie, Jue Wang, Ying Shan:
Learning Anchor Transformations for 3D Garment Animation. 491-500
- Hongwen Zhang , Siyou Lin, Ruizhi Shao, Yuxiang Zhang, Zerong Zheng, Han Huang, Yandong Guo, Yebin Liu:
CloSET: Modeling Clothed Humans on Continuous Surface with Explicit Template Decomposition. 501-511
- Yuliang Xiu, Jinlong Yang, Xu Cao, Dimitrios Tzionas , Michael J. Black:
ECON: Explicit Clothed humans Optimized via Normal integration. 512-523
- Chung-Yi Weng, Pratul P. Srinivasan, Brian Curless, Ira Kemelmacher-Shlizerman:
PersonNeRF : Personalized Reconstruction from Photo Collections. 524-533
- Xiaoxuan Ma, Jiajun Su, Chunyu Wang, Wentao Zhu, Yizhou Wang:
3D Human Mesh Estimation from Virtual Markers. 534-543
- Ziwei Yu, Chen Li, Linlin Yang, Xiaoxu Zheng, Michael Bi Mi, Gim Hee Lee, Angela Yao:
Overcoming the TradeOff between Accuracy and Plausibility in 3D Hand Shape Reconstruction. 544-553
- Yeonguk Oh, JoonKyu Park, Jaeha Kim, Gyeongsik Moon, Kyoung Mu Lee:
Recovering 3D Hand Mesh Sequence from a Single Blurry Image: A New Dataset and Temporal Unfolding. 554-563
- Congyi Wang, Feida Zhu, Shilei Wen:
MeMaHand: Exploiting Mesh-Mano Interaction for Single Image Two-Hand Reconstruction. 564-573
- Karthik Shetty, Annette Birkhold, Srikrishna Jaganathan, Norbert Strobel, Markus Kowarschik, Andreas K. Maier, Bernhard Egger :
PLIKS: A Pseudo-Linear Inverse Kinematic Solver for 3D Human Body Estimation. 574-584
- Juntian Zheng, Qingyuan Zheng, Lixing Fang, Yun Liu, Li Yi:
CAMS: CAnonicalized Manipulation Spaces for Category-Level Functional Hand-Object Manipulation Synthesis. 585-594
- Yuheng Jiang, Kaixin Yao, Zhuo Su, Zhehao Shen, Haimin Luo, Lan Xu:
Instant-NVR: Instant Neural Volumetric Rendering for Human-object Interactions from Monocular RGBD Stream. 595-605
- Bowen Wen, Jonathan Tremblay, Valts Blukis, Stephen Tyree, Thomas Müller, Alex Evans, Dieter Fox, Jan Kautz, Stan Birchfield:
BundleSDF: Neural 6-DoF Tracking and 3D Reconstruction of Unknown Objects. 606-617
- Xuan Ju, Ailing Zeng, Jianne Wang, Qiang Xu, Lei Zhang:
Human-Art: A Versatile Human-Centric Dataset Bridging Natural and Artificial Scenes. 618-629
- Mohammed Suhail, Erika Lu, Zhengqi Li, Noah Snavely, Leonid Sigal, Forrester Cole:
Omnimatte3D: Associating Objects and Their Effects in Unconstrained Monocular Video. 630-639
- Jathushan Rajasegaran, Georgios Pavlakos, Angjoo Kanazawa, Christoph Feichtenhofer, Jitendra Malik:
On the Benefits of 3D Pose and Tracking for Human Action Recognition. 640-649
- Li'an Zhuo, Jian Cao, Qi Wang, Bang Zhang, Liefeng Bo:
Towards Stable Human Pose Estimation via Cross-View Fusion and Foot Stabilization. 650-659
- Zigang Geng, Chunyu Wang, Yixuan Wei, Ze Liu, Houqiang Li, Han Hu:
Human Pose as Compositional Tokens. 660-671
- Qihao Liu, Adam Kortylewski, Alan L. Yuille:
PoseExaminer: Automated Testing of Out-of-Distribution Robustness in Human Pose and Shape Estimation. 672-681
- Yudi Dai , Yitai Lin, Xiping Lin, Chenglu Wen, Lan Xu, Hongwei Yi, Siqi Shen, Yuexin Ma, Cheng Wang:
SLOPER4D: A Scene-Aware Dataset for Global 4D Human Pose Estimation in Urban Environments. 682-692
- Linzhi Huang, Yulong Li, Hongbo Tian, Yue Yang, Xiangang Li, Weihong Deng, Jieping Ye:
Semi-Supervised 2D Human Pose Estimation Driven by Position Inconsistency Pseudo Label Correction Module. 693-703
- Sohyun Lee, Jaesung Rim, Boseung Jeong, Geonu Kim, Byungju Woo, Haechan Lee, Sunghyun Cho, Suha Kwak:
Human Pose Estimation in Extremely Low-Light Conditions. 704-714



- Riqiang Gao, Bin Lou , Zhoubing Xu, Dorin Comaniciu, Ali Kamen:
Flexible-C^m GAN: Towards Precise 3D Dose Prediction in Radiotherapy. 715-725
- Antyanta Bangunharcana, Ahmed Magd, Kyung-Soo Kim:
DualRefine: Self-Supervised Depth and Pose Estimation Through Iterative Epipolar Sampling and Refinement Toward Equilibrium. 726-738
- Yijia He, Bo Xu, Zhanpeng Ouyang, Hongdong Li:
A Rotation-Translation-Decoupled Solution for Robust and Efficient Visual-Inertial Initialization. 739-748
- Linus Härenstam-Nielsen, Niclas Zeller, Daniel Cremers:
Semidefinite Relaxations for Robust Multiview Triangulation. 749-757
- Zheheng Jiang, Hossein Rahmani, Sue Black, Bryan M. Williams:
A Probabilistic Attention Model with Occlusion-aware Texture Regression for 3D Hand Reconstruction from a Single RGB Image. 758-768
- Timo Bolkart, Tianye Li, Michael J. Black:
Instant Multi-View Head Capture through Learnable Registration. 768-779
- HyunJun Jung, Patrick Ruhkamp, Guangyao Zhai, Nikolas Brasch, Yitong Li, Yannick Verdie, Jifei Song, Yiren Zhou, Anil Armagan, Slobodan Ilic, Ales Leonardis, Nassir Navab, Benjamin Busam:
On the Importance of Accurate Geometry Data for Dense 3D Vision Tasks. 780-791
- Yinyu Nie, Angela Dai, Xiaoguang Han, Matthias Nießner:
Learning 3D Scene Priors with 2D Supervision. 792-802
- Tong Wu, Jiarui Zhang, Xiao Fu , Yuxin Wang , Jiawei Ren, Liang Pan, Wayne Wu, Lei Yang, Jiaqi Wang, Chen Qian, Dahua Lin, Ziwei Liu:
OmniObject3D: Large-Vocabulary 3D Object Dataset for Realistic Perception, Reconstruction and Generation. 803-814
- Songyou Peng, Kyle Genova, Chiyu Max Jiang, Andrea Tagliasacchi, Marc Pollefeys, Thomas A. Funkhouser:
OpenScene: 3D Scene Understanding with Open Vocabularies. 815-824
- Xu Cao, Hiroaki Santo, Fumio Okura, Yasuyuki Matsushita :
Multi-View Azimuth Stereo via Tangent Space Consistency. 825-834
- Yi-Ting Shen, Hyungtae Lee, Heesung Kwon, Shuvra S. Bhattacharyya:
Progressive Transformation Learning for Leveraging Virtual Images in Training. 835-844
- Yuanwen Yue, Theodora Kontogianni, Konrad Schindler, Francis Engelmann:
Connecting the Dots: Floorplan Reconstruction Using Two-Level Queries. 845-854
- Fabio Tosi, Alessio Tonioni, Daniele De Gregorio, Matteo Poggi:
NeRF-Supervised Deep Stereo. 855-866
- Fengyun Wang, Dong Zhang, Hanwang Zhang, Jinhui Tang, Qianru Sun:
Semantic Scene Completion with Cleaner Self. 867-877
- Haozheng Yu, Lu He, Bing Jian, Weiwei Feng, Shan Liu:
PanelNet: Understanding 360 Indoor Environment via Panel Representation. 878-887
- Avinash Paliwal, Andrii Tsarov, Nima Khademi Kalantari:
Implicit View-Time Interpolation of Stereo Videos Using Multi-Plane Disparities and Non-Uniform Coordinates. 888-898
- Wenjie Chang, Yueyi Zhang, Zhiwei Xiong:
Depth Estimation from Indoor Panoramas with Neural Scene Representation. 899-908
- Zehan Zheng, Danni Wu, Ruisi Lu, Fan Lu, Guang Chen, Changjun Jiang:
NeuralPCI: Spatio-Temporal Neural Field for 3D Point Cloud Multi-Frame Non-Linear Interpolation. 909-918
- Changjiang Cai , Pan Ji, Qingan Yan, Yi Xu:
RIAV-MVS: Recurrent-Indexing an Asymmetric Volume for Multi-View Stereo. 919-928
- Shitao Tang, Sicong Tang, Andrea Tagliasacchi, Ping Tan, Yasutaka Furukawa:
NeuMap: Neural Coordinate Mapping by Auto-Transdecoder for Camera Localization. 929-939
- Antoine Guédon, Tom Monnier, Pascal Monasse, Vincent Lepetit:
MACARONS: Mapping and Coverage Anticipation with RGB Online Self-Supervision. 940-951
- Xin Kong, Shikun Liu, Marwan Taher, Andrew J. Davison:
vMAP: Vectorised Object Mapping for Neural Field SLAM. 952-961
- Yunzhi Zhang, Shangzhe Wu, Noah Snavely, Jiajun Wu:
Seeing a Rose in Five Thousand Ways. 962-971
- Yihao Wang, Zhigang Wang, Bin Zhao, Dong Wang, Mulin Chen, Xuelong Li:
Propagate and Calibrate: Real-Time Passive Non-Line-of-Sight Tracking. 972-981
- Praneeth Chakravarthula, Jim Aldon D'Souza, Ethan Tseng, Joe Bartusek, Felix Heide:
Seeing With Sound: Long-Range Acoustic Beamforming for Multimodal Scene Understanding. 982-991
- Jia Zeng, Li Chen, Hanming Deng, Lewei Lu, Junchi Yan, Yu Qiao, Hongyang Li:
Distilling Focal Knowledge from Imperfect Expert for 3D Object Detection. 992-1001



- Zechuan Li, Hongshan Yu, Zhengeng Yang, Tom Tongjia Chen, Naveed Akhtar:
AShapeFormer : Semantics-Guided Object-Level Active Shape Encoding for 3D Object Detection via Transformers. 1012-1021
- Yinpeng Dong, Caixin Kang, Jinlai Zhang, Zijian Zhu, Yikai Wang, Xiao Yang, Hang Su, Xingxing Wei, Jun Zhu:
Benchmarking Robustness of 3D Object Detection to Common Corruptions in Autonomous Driving. 1022-1032
- Hang Xu, Xinyuan Liu, Qiang Zhao, Yike Ma, Chenggang Yan, Feng Dai:
Gaussian Label Distribution Learning for Spherical Image Object Detection. 1033-1042
- Ukcheol Shin , Jinsun Park, In So Kweon:
Deep Depth Estimation from Thermal Image. 1043-1053
- Chuanfu Shen, Fan Chao, Wei Wu, Rui Wang, George Q. Huang, Shiqi Yu:
LidarGait: Benchmarking 3D Gait Recognition with Point Clouds. 1054-1063
- Kunyu Wang, Xueyang Fu, Yukun Huang, Chengzhi Cao, Gege Shi , Zheng-Jun Zha:
Generalized UAV Object Detection via Frequency Domain Disentanglement. 1064-1073
- Yuwen Xiong, Wei-Chiu Ma, Jingkang Wang, Raquel Urtasun:
Learning Compact Representations for LiDAR Completion and Generation. 1074-1083
- Tian-Xing Xu, Yuan-Chen Guo, Yu-Kun Lai, Song-Hai Zhang:
CXTrack: Improving 3D Point Cloud Tracking with Contextual Information. 1084-1093
- Wei Ji, Jingjing Li, Cheng Bian, Zongwei Zhou, Jiaying Zhao, Alan L. Yuille, Li Cheng:
Multispectral Video Semantic Segmentation: A Benchmark Dataset and Baseline. 1094-1104
- Tao Lu, Xiang Ding, Haisong Liu, Gangshan Wu, Limin Wang:
LinK: Linear Kernel for LiDAR-based 3D Perception. 1105-1115
- Tarasha Khurana, Peiyun Hu, David Held, Deva Ramanan:
Point Cloud Forecasting as a Proxy for 4D Occupancy Forecasting. 1116-1124
- Ziyue Zhu, Qiang Meng, Xiao Wang, Ke Wang, Liujiang Yan, Jian Yang:
Curricular Object Manipulation in LiDAR-based Object Detection. 1125-1135
- Jiaming Zhang, Ruiping Liu, Hao Shi , Kailun Yang, Simon Reiß , Kunyu Peng, Haodong Fu, Kaiwei Wang, Rainer Stiefelhagen:
Delivering Arbitrary-Modal Semantic Segmentation. 1136-1147
- Haobo Jiang, Zheng Dang, Zhen Wei, Jin Xie, Jian Yang, Mathieu Salzmann:
Robust Outlier Rejection for 3D Registration with Variational Bayes. 1148-1157
- Zhenzhen Weng , Alexander S. Gorban, Jingwei Ji, Mahyar Najibi, Yin Zhou, Dragomir Anguelov:
3D Human Keypoints Estimation from Point Clouds in the Wild without Human Labels. 1158-1167
- Li Jiang , Zetong Yang, Shaoshuai Shi, Vladislav Golyanik, Dengxin Dai, Bernt Schiele:
Self-Supervised Pre-Training with Masked Shape Prediction for 3D Scene Understanding. 1168-1178
- Le Xue , Mingfei Gao, Chen Xing, Roberto Martín-Martín, Jiajun Wu, Caiming Xiong, Ran Xu, Juan Carlos Niebles, Silvio Savarese:
ULIP: Learning a Unified Representation of Language, Images, and Point Clouds for 3D Understanding. 1179-1189
- Yuheng Lu, Chenfeng Xu, Xiaobao Wei, Xiaodong Xie, Masayoshi Tomizuka, Kurt Keutzer, Shanghang Zhang:
Open-Vocabulary Point-Cloud Object Detection without 3D Annotation. 1190-1199
- Zhijian Liu, Xinyu Yang, Haotian Tang, Shang Yang, Song Han:
FlatFormer: Flattened Window Attention for Efficient Point Cloud Transformer. 1200-1211
- Zhiqiang Shen, Xiaoxiao Sheng, Longguang Wang, Yulan Guo, Qiong Liu, Xi Zhou:
PointCMP: Contrastive Mask Prediction for Self-supervised Learning on Point Cloud Videos. 1212-1222
- Minghan Zhu, Maani Ghaffari, William A. Clark, Huei Peng:
E2PN: Efficient SE(3)-Equivariant Point Network. 1223-1232
- Tao Xie, Shiguang Wang, Ke Wang, Linqi Yang, Zhiqiang Jiang, Xingcheng Zhang, Kun Dai, Rui Feng Li, Jian Cheng:
Poly-PC: A Polyhedral Network for Multiple Point Cloud Tasks at Once. 1233-1243
- Nan Zhang, Zhiyi Pan, Thomas H. Li, Wei Gao , Ge Li:
Improving Graph Representation for Point Cloud Segmentation via Attentive Filtering. 1244-1254
- Sheng Ao, Qingyong Hu, Hanyun Wang , Kai Xu, Yulan Guo:
BUFFER: Balancing Accuracy, Efficiency, and Generalizability in Point Cloud Registration. 1255-1264
- Bingnan Yang, Mi Zhang, Zhan Zhang, Zhili Zhang, Xiangyun Hu:
TopDiG: Class-agnostic Topological Directional Graph Extraction from Remote Sensing Images. 1265-1274
- Daniel Widdowson, Vitaliy Kurlin:
Recognizing Rigid Patterns of Unlabeled Point Clouds by Complete and Continuous Isometry Invariants with no False Negatives and no False Positives. 1275-1284
- Xu Zheng, Jinjing Zhu, Yexin Liu, Zidong Cao, Chong Fu, Lin Wang :
Both Style and Distortion Matter: Dual-Path Unsupervised Domain Adaptation for Panoramic Semantic Segmentation. 1285-1295
- Harshil Bhatia, Edith Treitschke, Zorah Lähner , Marcel Seelbach Benkner, Michael Moeller, Christian Theobalt, Vladislav Golyanik:

**CCuantuMM: Cycle-Consistent Quantum-Hybrid Matching of Multiple Shapes.** 1296-1305

Guilherme A. Potje, Felipe Cadar, André Araújo, Renato Martins, Erickson R. Nascimento:
Enhancing Deformable Local Features by Jointly Learning to Detect and Describe Keypoints. 1306-1315

Souhaib Attaiki, Maks Ovsjanikov:
Understanding and Improving Features Learned in Deep Functional Maps. 1316-1326

Haoliang Zhao, Huizhou Zhou, Yongjun Zhang, Jie Chen, Yitong Yang, Yong Zhao:
High-Frequency Stereo Matching Network. 1327-1336

Qiaole Dong , Chenjie Cao, Yanwei Fu :
Rethinking Optical Flow from Geometric Matching Consistent Perspective. 1337-1347

Shun Fang, Zhengqin Xu, Shiqian Wu, Shoulie Xie:
Efficient Robust Principal Component Analysis via Block Krylov Iteration and CUR Decomposition. 1348-1357

Bingchen Yang, Haiyong Jiang, Hao Pan, Jun Xiao:
VectorFloorSeg: Two-Stream Graph Attention Network for Vectorized Roughcast Floorplan Segmentation. 1358-1367

Shaoheng Fang, Zi Wang, Yiqi Zhong, Junhao Ge, Siheng Chen:
TBP-Former: Learning Temporal Bird's-Eye-View Pyramid for Joint Perception and Prediction in Vision-Centric Autonomous Driving. 1368-1378

Ben Agro, Quinlan Sykora, Sergio Casas, Raquel Urtasun:
Implicit Occupancy Flow Fields for Perception and Prediction in Self-Driving. 1379-1388

Ze Yang, Yun Chen, Jingkang Wang, Sivabalan Manivasagam, Wei-Chiu Ma, Anqi Joyce Yang, Raquel Urtasun:
UniSim: A Neural Closed-Loop Sensor Simulator. 1389-1399

Yuning Wang, Pu Zhang, Lei Bai , Jianru Xue:
FEND: A Future Enhanced Distribution-Aware Contrastive Learning Framework for Long-Tail Trajectory Prediction. 1400-1409

Chenxin Xu, Robby T. Tan, Yuhong Tan, Siheng Chen, Yu Guang Wang, Xinchao Wang, Yanfeng Wang:
EqMotion: Equivariant Multi-Agent Motion Prediction with Invariant Interaction Reasoning. 1410-1420

Guoqiang Zhang, Kenta Niwa, W. Bastiaan Kleijn :
Lookahead Diffusion Probabilistic Models for Refining Mean Estimation. 1421-1429

Ruihan Yang, Ge Yang, Xiaolong Wang:
Neural Volumetric Memory for Visual Locomotion Control. 1430-1440

Sounak Mondal, Zhibo Yang, Seoyoung Ahn, Dimitris Samaras, Gregory J. Zelinsky, Minh Hoai:
Gazeformer: Scalable, Effective and Fast Prediction of Goal-Directed Human Attention. 1441-1450

Luca De Luigi, Ren Li, Benoît Guillard, Mathieu Salzmann, Pascal Fua:
DrapeNet: Garment Generation and Self-Supervised Draping. 1451-1460

Mingzhen Huang, Xiaoxing Li, Jun Hu, Honghong Peng, Siwei Lyu:
Tracking Multiple Deformable Objects in Egocentric Videos. 1461-1471

Zhengwei Yang , Meng Lin, Xian Zhong, Yu Wu, Zheng Wang:
Good is Bad: Causality Inspired Cloth-debiasing for Cloth-changing Person Re-identification. 1472-1481

Xuan-Bac Nguyen, Chi Nhan Duong, Xin Li, Susan Gauch, Han-Seok Seo, Khoa Luu:
Micron-BERT: BERT-Based Facial Micro-Expression Recognition. 1482-1492

Zhixi Cai , Shreya Ghosh , Kalin Stefanov, Abhinav Dhall, Jianfei Cai, Hamid Rezatofighi, Reza Haffari, Munawar Hayat:
MARLIN: Masked Autoencoder for facial video Representation LearnINg. 1493-1504

Jiazhi Guan, Zhanwang Zhang, Hang Zhou, Tianshu Hu, Kaisiyuan Wang, Dongliang He, Haocheng Feng, Jingtuo Liu, Errui Ding, Ziwei Liu, Jingdong Wang:
StyleSync: High-Fidelity Generalized and Personalized Lip Sync in Style-Based Generator. 1505-1515

Samuel Clarke, Ruohan Gao, Mason L. Wang, Mark Rau, Julia Xu, Jui-Hsien Wang, Doug L. James , Jiajun Wu:
REALIMPACT: A Dataset of Impact Sound Fields for Real Objects. 1516-1525

Xiaoyu Zhu, Po-Yao Huang, Junwei Liang , Celso M. de Melo, Alexander G. Hauptmann:
STMT: A Spatial-Temporal Mesh Transformer for MoCap-Based Action Recognition. 1526-1536

Xueyan Huang, Yueyi Zhang, Zhiwei Xiong:
Progressive Spatio-temporal Alignment for Efficient Event-based Motion Estimation. 1537-1546

Manasi Muglikar, Leonard Bauersfeld, Diederik Paul Moeys, Davide Scaramuzza :
Event-Based Shape from Polarization. 1547-1556

Yunfan Lu, Zipeng Wang, Minjie Liu, Hongjian Wang, Lin Wang :
Learning Spatial-Temporal Implicit Neural Representations for Event-Guided Video Super-Resolution. 1557-1567

Junheum Park, Jintae Kim , Chang-Su Kim:
BiFormer: Learning Bilateral Motion Estimation via Bilateral Transformer for 4K Video Frame Interpolation. 1568-1577

Xin Jin, Longhai Wu, Jie Chen, Youxin Chen, Jayoon Koo, Cheul-Hee Hahm:
A Unified Pyramid Recurrent Network for Video Frame Interpolation. 1578-1587



- Wenming Weng, Yueyi Zhang, Zhiwei Xiong:
Event-based Blurry Frame Interpolation under Blind Exposure. 1588-1598
- Xiaoyu Shi, Zhaoyang Huang, Dasong Li, Manyuan Zhang, Ka Chun Cheung, Simon See, Hongwei Qin, Jifeng Dai, Hongsheng Li :
FlowFormer++: Masked Cost Volume Autoencoding for Pretraining Optical Flow Estimation. 1599-1610
- Ce Zheng, Xianpeng Liu, Guo-Jun Qi, Chen Chen:
POTTER: Pooling Attention Transformer for Efficient Human Mesh Recovery. 1611-1620
- Yuesong Wang, Zhaojie Zeng, Tao Guan, Wei Yang, Zhuo Chen, Wenkai Liu, Luoyuan Xu, Yawei Luo:
Adaptive Patch Deformation for Textureless-Resilient Multi-View Stereo. 1621-1630
- Zhenjie Yu, Shuang Li, Yirui Shen , Chi Harold Liu, Shuigen Wang:
On the Difficulty of Unpaired Infrared-to-Visible Video Translation: Fine-Grained Content-Rich Patches Transfer. 1631-1640
- Aniket Dashputre , Vishwanath Saragadam, Emma Alexander, Florian Willomitzer, Aggelos K. Katsaggelos, Ashok Veeraraghavan, Oliver Cossairt:
Thermal Spread Functions (TSF): Physics-Guided Material Classification. 1641-1650
- Xuhai Chen, Jiangning Zhang, Chao Xu, Yabiao Wang, Chengjie Wang, Yong Liu:
Better "CMOS" Produces Clearer Images: Learning Space-Variant Blur Estimation for Blind Image Super-Resolution. 1651-1661
- Yuhui Wu, Chen Pan, Guoqing Wang, Yang Yang, Jiwei Wei, Chongyi Li, Heng Tao Shen:
Learning Semantic-Aware Knowledge Guidance for Low-Light Image Enhancement. 1662-1671
- Zeyu Xiao, Yutong Liu, Ruisheng Gao, Zhiwei Xiong:
CutMIB: Boosting Light Field Super-Resolution via Multi-View Image Blending. 1672-1682
- Zixuan Fu, Lanqing Guo, Bihan Wen :
sRGB Real Noise Synthesizing with Neighboring Correlation-Aware Noise Model. 1683-1691
- Haoyu Chen, Jinjin Gu, Yihao Liu , Salma Abdel Magid, Chao Dong, Qiong Wang, Hanspeter Pfister, Lei Zhu:
Masked Image Training for Generalizable Deep Image Denoising. 1692-1703
- Zhixin Wang, Ziying Zhang, Xiaoyun Zhang, Huangjie Zheng, Mingyuan Zhou, Ya Zhang, Yanfeng Wang:
DR2: Diffusion-Based Robust Degradation Remover for Blind Face Restoration. 1704-1713
- Xin Li, Bingchen Li, Xin Jin, Cuiling Lan, Zhibo Chen:
Learning Distortion Invariant Representation for Image Restoration from a Causality Perspective. 1714-1724
- Seung Ho Park , Young-Su Moon, Nam Ik Cho:
Perception-Oriented Single Image Super-Resolution using Optimal Objective Estimation. 1725-1735
- Xinmiao Lin, Yikang Li, Jenhao Hsiao, Chiuman Ho, Yu Kong:
Catch Missing Details: Image Reconstruction with Frequency Augmented Variational Autoencoder. 1736-1745
- Zicheng Zhang, Wei Wu , Wei Sun, Danyang Tu, Wei Lu, Xiongkuo Min, Ying Chen, Guangtao Zhai:
MD-VQA: Multi-Dimensional Quality Assessment for UGC Live Videos. 1746-1755
- Senmao Tian, Ming Lu, Jiaming Liu, Yandong Guo, Yurong Chen, Shunli Zhang:
CABM: Content-Aware Bit Mapping for Single Image Super-Resolution Network with Large Input. 1756-1765
- Ann-Christin Woerl, Jan Disselhoff , Michael Wand:
Initialization Noise in Image Gradients and Saliency Maps. 1766-1775
- Jie-En Yao, Li-Yuan Tsao, Yi-Chen Lo, Roy Tseng, Chia-Che Chang, Chun-Yi Lee:
Local Implicit Normalizing Flow for Arbitrary-Scale Image Super-Resolution. 1776-1785
- Xiaohang Wang, Xuanhong Chen, Bingbing Ni, Hang Wang, Zhengyan Tong, Yutian Liu:
Deep Arbitrary-Scale Image Super-Resolution via Scale-Equivariance Pursuit. 1786-1795
- Jiezhang Cao , Qin Wang, Yongqin Xian, Yawei Li , Bingbing Ni, Zhiming Pi, Kai Zhang, Yulun Zhang, Radu Timofte , Luc Van Gool:
CiaoSR: Continuous Implicit Attention-in-Attention Network for Arbitrary-Scale Image Super-Resolution. 1796-1807
- Yishun Dou, Zhong Zheng, Qiaoqiao Jin, Bingbing Ni:
Multiplicative Fourier Level of Detail. 1808-1817
- Ling Zhang, Yinghao He, Qing Zhang, Zheng Liu, Xiaolong Zhang, Chunxia Xiao:
Document Image Shadow Removal Guided by Color-Aware Background. 1818-1827
- Hamza Pehlivan, Yusuf Dalva, Aysegul Dundar:
StyleRes: Transforming the Residuals for Real Image Editing with StyleGAN. 1828-1837
- Sijie Zhu, Zhe Lin, Scott Cohen, Jason Kuen, Zhifei Zhang, Chen Chen:
TopNet: Transformer-Based Object Placement Network for Image Compositing. 1838-1847
- Zeqing Xia, Bojun Xiong, Zhouhui Lian:
VecFontSDF: Learning to Reconstruct and Synthesize High-Quality Vector Fonts via Signed Distance Functions. 1848-1857
- Chi Wang , Min Zhou, Tiezheng Ge, Yuning Jiang, Hujun Bao, Weiwei Xu:
CF-Font: Content Fusion for Few-Shot Font Generation. 1858-1867



- Wuyang Luo, Su Yang, Xinjian Zhang, Weishan Zhang :
SIEDOB: Semantic Image Editing by Disentangling Object and Background. 1868-1878
- Dina Bashkirova, José Lezama, Kihyuk Sohn, Kate Saenko, Irfan Essa:
MaskSketch: Unpaired Structure-guided Masked Image Generation. 1879-1889
- Inwoo Hwang, Hyeonwoo Kim, Young Min Kim:
Text2Scene: Text-driven Indoor Scene Stylistization with Part-Aware Details. 1890-1899
- Qiucheng Wu, Yujian Liu, Handong Zhao, Ajinkya Kale, Trung Bui, Tong Yu, Zhe Lin, Yang Zhang, Shiyu Chang:
Uncovering the Disentanglement Capability in Text-to-Image Diffusion Models. 1900-1910
- Ajay Jain, Amber Xie, Pieter Abbeel:
VectorFusion: Text-to-SVG by Abstracting Pixel-Based Diffusion Models. 1911-1920
- Narek Tumanyan, Michal Geyer, Shai Bagon, Tali Dekel:
Plug-and-Play Diffusion Features for Text-Driven Image-to-Image Translation. 1921-1930
- Nupur Kumari, Bingliang Zhang, Richard Zhang, Eli Shechtman, Jun-Yan Zhu:
Multi-Concept Customization of Text-to-Image Diffusion. 1931-1941
- Mude Hui, Zhizheng Zhang, Xiaoyi Zhang, Wenxuan Xie, Yuwang Wang, Yan Lu:
Unifying Layout Generation with a Decoupled Diffusion Model. 1942-1951
- Bo Li, Kaitao Xue , Bin Liu, Yu-Kun Lai:
BBDM: Image-to-Image Translation with Brownian Bridge Diffusion Models. 1952-1961
- Hyojun Go, Yunsung Lee, Jin Young Kim, Seunghyun Lee, Myeongho Jeong, Hyun Seung Lee, Seungtaek Choi:
Towards Practical Plug-and-Play Diffusion Models. 1962-1971
- Yuzhang Shang, Zhihang Yuan , Bin Xie, Bingzhe Wu , Yan Yan:
Post-Training Quantization on Diffusion Models. 1972-1981
- Shuai Shen, Wenliang Zhao, Zibin Meng, Wanhua Li, Zheng Zhu, Jie Zhou, Jiwen Lu:
DiffTalk: Crafting Diffusion Models for Generalized Audio-Driven Portraits Animation. 1982-1991
- Kwanyong Park, Sanghyun Woo, Seoung Wug Oh, In So Kweon, Joon-Young Lee:
Mask-Guided Matting in the Wild. 1992-2001
- Mengqi Huang, Zhendong Mao, Quan Wang, Yongdong Zhang:
Not All Image Regions Matter: Masked Vector Quantization for Autoregressive Image Generation. 2002-2011
- Yingwei Wang, Takashi Isobe, Xu Jia, Xin Tao, Huchuan Lu, Yu-Wing Tai:
Compression-Aware Video Super-Resolution. 2012-2021
- Nilesh A. Ahuja, Parual Datta, Bhavya Kanzariya, V. Srinivasa Somayazulu, Omesh Tickoo:
Neural Rate Estimator and Unsupervised Learning for Efficient Distributed Image Analytics in Split-DNN models. 2022-2030
- Qi Zhao, M. Salman Asif, Zhan Ma:
DNeRV: Modeling Inherent Dynamics via Difference Neural Representation for Videos. 2031-2040
- Rajhans Singh, Ankita Shukla, Pavan K. Turaga:
Polynomial Implicit Neural Representations For Large Diverse Datasets. 2041-2051
- Yutaro Shigeto, Masashi Shimbo, Yuya Yoshikawa, Akikazu Takeuchi:
Learning Decorrelated Representations Efficiently Using Fast Fourier Transform. 2052-2060
- Xuanyao Chen, Zhijian Liu, Haotian Tang, Li Yi, Hang Zhao, Song Han:
SparseViT: Revisiting Activation Sparsity for Efficient High-Resolution Vision Transformer. 2061-2070
- Haram Choi, Jeongmin Lee, Jihoon Yang:
N-Gram in Swin Transformers for Efficient Lightweight Image Super-Resolution. 2071-2081
- Xuran Pan, Tianzhu Ye, Zhuofan Xia , Shiji Song, Gao Huang:
Slide-Transformer: Hierarchical Vision Transformer with Local Self-Attention. 2082-2091
- Siyuan Wei, Tianzhu Ye, Shen Zhang, Yao Tang, Jiajun Liang:
Joint Token Pruning and Squeezing Towards More Aggressive Compression of Vision Transformers. 2092-2101
- Baifeng Shi, Trevor Darrell, Xin Wang:
Top-Down Visual Attention from Analysis by Synthesis. 2102-2112
- Markus Frey, Christian F. Doeller, Caswell Barry:
Probing Neural Representations of Scene Perception in a Hippocampally Dependent Task Using Artificial Neural Networks. 2113-2121
- Haoqing Wang, Yehui Tang, Yunhe Wang, Jianyuan Guo , Zhi-Hong Deng, Kai Han:
Masked Image Modeling with Local Multi-Scale Reconstruction. 2122-2131
- Chenxin Tao, Xizhou Zhu, Weijie Su, Gao Huang, Bin Li, Jie Zhou, Yu Qiao, Xiaogang Wang, Jifeng Dai:
Siamese Image Modeling for Self-Supervised Vision Representation Learning. 2132-2141
- Tianhong Li, Huiwen Chang, Shlok Kumar Mishra, Han Zhang, Dina Katahi, Dilip Krishnan:
MAGE: MAsked Generative Encoder to Unify Representation Learning and Image Synthesis. 2142-2152
- Yukang Zhang, Hanzi Wang:
Diverse Embedding Expansion Network and Low-Light Cross-Modality Benchmark for Visible-Infrared Person Re-



- identification. 2153-2162
- Suhang Ye, Yingyi Zhang, Jie Hu, Liujuan Cao, Shengchuan Zhang, Lei Shen, Jun Wang, Shouhong Ding, Rongrong Ji:
DistilPose: Tokenized Pose Regression with Heatmap Distillation. 2163-2172
- Hao Tang, Zhenyu Zhang, Humphrey Shi , Bo Li, Ling Shao, Nicu Sebe , Radu Timofte , Luc Van Gool:
Graph Transformer GANs for Graph-Constrained House Generation. 2173-2182
- Mang Tik Chiu, Xuaner Zhang, Zijun Wei, Yuqian Zhou, Eli Shechtman, Connelly Barnes, Zhe Lin, Florian Kainz, Sohrab Amirghodsi, Humphrey Shi
Automatic High Resolution Wire Segmentation and Removal. 2183-2192
- Adnan Firoze , Cameron Wingren, Raymond A. Yeh, Bedrich Benes, Daniel G. Aliaga:
Tree Instance Segmentation with Temporal Contour Graph. 2193-2202
- Jungin Park, Jiyoung Lee , Kwanghoon Sohn:
Dual-Path Adaptation from Image to Video Transformers. 2203-2213
- A. J. Piergiovanni, Weicheng Kuo, Anelia Angelova:
Rethinking Video ViTs: Sparse Video Tubes for Joint Image and Video Learning. 2214-2224
- Heng Zhang, Daqing Liu, Qi Zheng, Bing Su:
Modeling Video as Stochastic Processes for Fine-Grained Video Representation Learning. 2225-2234
- Xinyu Sun, Peihao Chen, Liangwei Chen, Changhao Li, Thomas H. Li, Mingkui Tan, Chuang Gan:
Masked Motion Encoding for Self-Supervised Video Representation Learning. 2235-2245
- Yurong Zhang, Liulei Li, Wenguan Wang , Rong Xie, Li Song, Wenjun Zhang:
Boosting Video Object Segmentation via Space-Time Correspondence Learning. 2246-2256
- Kun Yan, Xiao Li, Fangyun Wei, Jinglu Wang , Chenbin Zhang, Ping Wang, Yan Lu:
Two-shot Video Object Segmentation. 2257-2267
- Junke Wang, Dongdong Chen, Zuxuan Wu, Chong Luo, Chuanxin Tang, Xiyang Dai, Yucheng Zhao, Yujia Xie, Lu Yuan, Yu-Gang Jiang:
Look Before You Match: Instance Understanding Matters in Video Object Segmentation. 2268-2278
- Rui Li, Dong Liu :
Spatial-then-Temporal Self-Supervised Learning for Video Correspondence. 2279-2288
- Yogesh Kumar , Anand Mishra:
Few-Shot Referring Relationships in Videos. 2289-2298
- Yan-Bo Lin, Yi-Lin Sung, Jie Lei, Mohit Bansal, Gedas Bertasius:
Vision Transformers are Parameter-Efficient Audio-Visual Learners. 2299-2309
- Zihui Xue, Yale Song, Kristen Grauman, Lorenzo Torresani:
Egocentric Video Task Translation. 2310-2320
- Sicheng Yang, Zhiyong Wu, Minglei Li, Zhensong Zhang, Lei Hao, Weihong Bao, Haolin Zhuang:
QPGesture: Quantization-Based and Phase-Guided Motion Matching for Natural Speech-Driven Gesture Generation. 2321-2330
- Mingyang Sun, Mengchen Zhao, Yaqing Hou, Minglei Li, Huang Xu, Songcen Xu, Jianye Hao:
Co-speech Gesture Synthesis by Reinforcement Learning with Contrastive Pretrained Rewards. 2331-2340
- Ishan Rajendrakumar Dave, Mamshad Nayeem Rizve, Chen Chen, Mubarak Shah:
TimeBalance: Temporally-Invariant and Temporally-Distinctive Video Representations for Semi-Supervised Action Recognition. 2341-2352
- Xingyi Zhou, Anurag Arnab, Chen Sun, Cordelia Schmid:
How can objects help action recognition? 2353-2362
- Lilang Lin, Jiahang Zhang, Jiaying Liu:
Actionlet-Dependent Contrastive Learning for Unsupervised Skeleton-Based Action Recognition. 2363-2372
- Pilhyeon Lee, Taeoh Kim, Minho Shim, Dongyoong Wee, Hyeran Byun:
Decomposed Cross-Modal Distillation for RGB-based Temporal Action Detection. 2373-2383
- Beatrice van Amsterdam, Abdolrahim Kadkhodamohammadi, Imanol Luengo, Danail Stoyanov:
ASPnet: Action Segmentation with Shared-Private Representation of Multiple Data Sources. 2384-2393
- Huan Ren, Wenfei Yang, Tianzhu Zhang, Yongdong Zhang:
Proposal-Based Multiple Instance Learning for Weakly-Supervised Temporal Action Localization. 2394-2404
- Shiyi Zhang, Wenzun Dai, Sujia Wang, Xiangwei Shen, Jiwen Lu, Jie Zhou, Yansong Tang:
LOGO: A Long-Form Video Dataset for Group Action Quality Assessment. 2405-2414
- Toby Perrett, Saptarshi Sinha, Tilo Burghardt, Majid Mirmehdi, Dima Damen:
Use Your Head: Improving Long-Tail Video Recognition. 2415-2425
- Yuexi Du, Ziyang Chen, Justin Salamon, Bryan Russell, Andrew Owens:
Conditional Generation of Audio from Video via Foley Analogies. 2426-2436
- Sixun Dong, Huazhang Hu, Dongze Lian, Weixin Luo, Yicheng Qian, Shenghua Gao:
Weakly Supervised Video Representation Learning with Unaligned Text for Sequential Videos. 2437-2447



- Xiang Fang, Daizong Liu, Pan Zhou, Guoshun Nan:
You Can Ground Earlier than See: An Effective and Efficient Pipeline for Temporal Sentence Grounding in Compressed Videos. 2448-2460
- Paul Voigtlaender, Soravit Changpinyo, Jordi Pont-Tuset, Radu Soricut, Vittorio Ferrari:
Connecting Vision and Language with Video Localized Narratives. 2461-2471
- Peng Jin, Jinfan Huang, Pengfei Xiong, Shangxuan Tian, Chang Liu, Xiangyang Ji, Li Yuan, Jie Chen:
Video-Text as Game Players: Hierarchical Banzhaf Interaction for Cross-Modal Representation Learning. 2472-2482
- Jiahao Zhang , Anoop Cherian, Yanbin Liu, Yizhak Ben-Shabat, Cristian Rodriguez Opazo, Stephen Gould:
Aligning Step-by-Step Instructional Diagrams to Video Demonstrations. 2483-2492
- Tanzila Rahman, Hsin-Ying Lee, Jian Ren, Sergey Tulyakov, Shweta Mahajan, Leonid Sigal:
Make-A-Story: Visual Memory Conditioned Consistent Story Generation. 2493-2502
- Piyush Bagad, Makarand Tapaswi, Cees G. M. Snoek:
Test of Time: Instilling Video-Language Models with a Sense of Time. 2503-2516
- Dhruv Srivastava , Aditya Kumar Singh, Makarand Tapaswi:
How You Feelin'? Learning Emotions and Mental States in Movie Scenes. 2517-2528
- Lianyu Hu, Liqing Gao, Zekang Liu, Wei Feng:
Continuous Sign Language Recognition with Correlation Network. 2529-2539
- Changsong Wen, Guoli Jia, Jufeng Yang:
DIP: Dual Incongruity Perceiving Network for Sarcasm Detection. 2540-2550
- Aoxiong Yin, Tianyun Zhong, Li Tang, Weike Jin, Tao Jin, Zhou Zhao:
Gloss Attention for Gloss-free Sign Language Translation. 2551-2562
- Heming Du, Lincheng Li, Zi Huang , Xin Yu :
Object-Goal Visual Navigation via Effective Exploration of Relations Among Historical Navigation States. 2563-2573
- Zijiao Yang , Arjun Majumdar, Stefan Lee:
Behavioral Analysis of Vision-and-Language Navigation Agents. 2574-2582
- Xiangyang Li, Zihan Wang, Jiahao Yang, Yaowei Wang, Shuqiang Jiang:
KERM: Knowledge Enhanced Reasoning for Vision-and-Language Navigation. 2583-2592
- Mengmeng Xu, Yanghao Li, Cheng-Yang Fu, Bernard Ghanem , Tao Xiang, Juan-Manuel Pérez-Rúa:
Where is my Wallet? Modeling Object Proposal Sets for Egocentric Visual Query Localization. 2593-2603
- Yaowei Li, Ruijie Quan, Linchao Zhu, Yi Yang:
Efficient Multimodal Fusion via Interactive Prompting. 2604-2613
- Joy Hsu, Jiayuan Mao, Jiajun Wu:
NS3D: Neuro-Symbolic Grounding of 3D Objects and Relations. 2614-2623
- Burak Uzkent, Amanmeet Garg, Wentao Zhu, Keval Doshi, Jingru Yi, Xiaolong Wang, Mohamed Omar:
Dynamic Inference with Grounding Based Vision and Language Models. 2624-2633
- Shuquan Ye , Yujia Xie, Dongdong Chen, Yichong Xu, Lu Yuan, Chenguang Zhu, Jing Liao :
Improving Commonsense in Vision-Language Models via Knowledge Graph Riddles. 2634-2645
- Wei Suo, Mengyang Sun, Weisong Liu, Yiqi Gao, Peng Wang, Yanning Zhang, Qi Wu:
S³C: Semi-Supervised VQA Natural Language Explanation via Self-Critical Learning. 2646-2656
- Sivan Doveh, Assaf Arbelle, Sivan Harary, Eli Schwartz, Roei Herzig, Raja Giryes, Rogério Feris, Rameswar Panda, Shimon Ullman, Leonid Karlinsky:
Teaching Structured Vision & Language Concepts to Vision & Language Models. 2657-2668
- Xiao Han, Xiatian Zhu, Licheng Yu, Li Zhang, Yi-Zhe Song, Tao Xiang:
FAME-ViL: Multi-Tasking Vision-Language Model for Heterogeneous Fashion Tasks. 2669-2680
- Hao Li, Jinguo Zhu, Xiaohu Jiang, Xizhou Zhu, Hongsheng Li , Chun Yuan, Xiaohua Wang, Yu Qiao, Xiaogang Wang, Wenhui Wang, Jifeng Dai:
Uni-Perceiver v2: A Generalist Model for Large-Scale Vision and Vision-Language Tasks. 2691-2700
- Shi Chen , Nachiappan Valliappan, Shaolei Shen, Xinyu Ye, Kai Kohlhoff, Junfeng He:
Learning from Unique Perspectives: User-aware Saliency Modeling. 2701-2710
- Thomas Fel, Agustin Martin Picard, Louis Béthune, Thibaut Boissin, David Vigouroux, Julien Colin, Rémi Cadène, Thomas Serre:
CRAFT: Concept Recursive Activation Factorization for Explainability. 2711-2721
- Chengzhi Mao, Revant Teotia, Amrutha Sundar, Sachit Menon, Junfeng Yang, Xin Wang, Carl Vondrick:
Doubly Right Object Recognition: A Why Prompt for Visual Rationales. 2722-2732
- Ayan Kumar Bhunia, Subhadeep Koley, Amandeep Kumar, Aneeshan Sain , Pinaki Nath Chowdhury, Tao Xiang, Yi-Zhe Song :
Sketch2Saliency: Learning to Detect Salient Objects from Human Drawings. 2733-2743
- Meike Nauta, Jörg Schlötterer, Maurice van Keulen, Christin Seifert:
PIP-Net: Patch-Based Intuitive Prototypes for Interpretable Image Classification. 2744-2753



- Aneeshan Sain , Ayan Kumar Bhunia, Pinaki Nath Chowdhury, Subhadeep Koley, Tao Xiang, Yi-Zhe Song :
CLIP for All Things Zero-Shot Sketch-Based Image Retrieval, Fine-Grained or Not. 2765-2775
- Yixuan Wei, Yue Cao, Zheng Zhang, Houwen Peng, Zhiliang Yao, Zhenda Xie, Han Hu, Baining Guo:
iCLIP: Bridging Image Classification and Contrastive Language-Image Pre-training for Visual Recognition. 2776-2786
- Ding Jiang, Mang Ye :
Cross-Modal Implicit Relation Reasoning and Aligning for Text-to-Image Person Retrieval. 2787-2797
- Jaeyoo Park, Bohyung Han:
Multi-Modal Representation Learning with Text-Driven Soft Masks. 2798-2807
- Zixian Guo, Bowen Dong, Zhilong Ji, Jinfeng Bai, Yiwen Guo, Wangmeng Zuo:
Texts as Images in Prompt Tuning for Multi-Label Image Recognition. 2808-2817
- Mehdi Cherti, Romain Beaumont, Ross Wightman, Mitchell Wortsman, Gabriel Ilharco, Cade Gordon, Christoph Schuhmann, Ludwig Schmidt, Jenia Jitsev:
Reproducible Scaling Laws for Contrastive Language-Image Learning. 2818-2829
- Zheng Wang, Zhenwei Gao, Kangshuai Guo, Yang Yang, Xiaoming Wang, Heng Tao Shen:
Multilateral Semantic Relations Modeling for Image Text Retrieval. 2830-2839
- Rita Ramos, Bruno Martins, Desmond Elliott , Yova Kementchedjhieva:
Smallcap: Lightweight Image Captioning Prompted with Retrieval Augmentation. 2840-2849
- Tinglei Feng, Jiaxuan Liu, Jufeng Yang:
Probing Sentiment-Oriented PreTraining Inspired by Human Sentiment Perception Mechanism. 2850-2860
- Kuniaki Saito, Kihyuk Sohn, Xiang Zhang, Chun-Liang Li, Chen-Yu Lee, Kate Saenko, Tomas Pfister:
Prefix Conditioning Unifies Language and Label Supervision. 2861-2870
- Yuchen Ren, Zhendong Mao, Shancheng Fang, Yan Lu, Tong He, Hao Du, Yongdong Zhang, Wanli Ouyang:
Crossing the Gap: Domain Generalization for Image Captioning. 2871-2880
- Weijie Tu, Weijian Deng, Tom Gedeon, Liang Zheng:
A Bag-of-Prototypes Representation for Dataset-Level Applications. 2881-2892
- Dingkang Liang, Jiahao Xie, Zhiqiang Zou, Xiaoqing Ye, Wei Xu, Xiang Bai:
CrowdCLIP: Unsupervised Crowd Counting via Vision-Language Model. 2893-2903
- Jianfeng He, Yuan Gao , Tianzhu Zhang, Zhe Zhang, Feng Wu:
D²Former: Jointly Learning Hierarchical Detectors and Contextual Descriptors via Agent-Based Transformers. 2904-2914
- Yong Zhang, Yingwei Pan, Ting Yao, Rui Huang, Tao Mei, Chang Wen Chen:
Learning to Generate Language-Supervised and Open-Vocabulary Scene Graph Using Pre-Trained Visual-Semantic Space. 2915-2924
- Sanghyun Kim, Deunsol Jung, Minsu Cho:
Relational Context Learning for Human-Object Interaction Detection. 2925-2934
- Jilan Xu, Junlin Hou, Yuejie Zhang, Rui Feng, Yi Wang, Yu Qiao, Weidi Xie:
Learning Open-Vocabulary Semantic Segmentation Models From Natural Language Supervision. 2935-2944
- Mengde Xu, Zheng Zhang, Fangyun Wei, Han Hu, Xiang Bai:
Side Adapter Network for Open-Vocabulary Semantic Segmentation. 2945-2954
- Jiarui Xu, Sifei Liu, Arash Vahdat, Wonmin Byeon, Xiaolong Wang, Shalini De Mello:
Open-Vocabulary Panoptic Segmentation with Text-to-Image Diffusion Models. 2955-2966
- Sukmin Yun, Seong Hyeon Park , Paul Hongseok Seo, Jinwoo Shin:
IFSeg: Image-free Semantic Segmentation via Vision-Language Model. 2967-2977
- Haoran Geng, Ziming Li, Yiran Geng, Jiayi Chen, Hao Dong, He Wang:
PartManip: Learning Cross-Category Generalizable Part Manipulation Policy from Point Cloud Observations. 2978-2988
- Jitesh Jain, Jiachen Li, MangTik Chiu, Ali Hassani, Nikita Orlov, Humphrey Shi :
OneFormer: One Transformer to Rule Universal Image Segmentation. 2989-2998
- Xinyu Liu, Beiwen Tian, Zhen Wang, Rui Wang, Kehua Sheng, Bo Zhang, Hao Zhao, Guyue Zhou:
Delving into Shape-aware Zero-shot Semantic Segmentation. 2999-3009
- Fabio Cermelli, Matthieu Cord, Arthur Douillard:
CoFormer: Continual Learning in Semantic and Panoptic Segmentation. 3010-3020
- Mengxue Qu, Yu Wu, Yunchao Wei, Wu Liu, Xiaodan Liang, Yao Zhao:
Learning to Segment Every Referring Object Point by Point. 3021-3030
- Zhizheng Liu, Francesco Milano, Jonas Frey, Roland Siegwart, Hermann Blum, Cesar Cadena:
Unsupervised Continual Semantic Adaptation Through Neural Rendering. 3031-3040
- Feng Li, Hao Zhang, Huaizhe Xu, Shilong Liu, Lei Zhang, Lionel M. Ni, Heung-Yeung Shum:
Mask DINO: Towards A Unified Transformer-based Framework for Object Detection and Segmentation. 3041-3050



- Hengcan Shi, Munawar Hayat, Jianfei Cai:
Transformer Scale Gate for Semantic Segmentation. 3051-3060
- Wei Huang, Chang Chen, Yong Li, Jiacheng Li, Cheng Li, Fenglong Song, Youliang Yan, Zhiwei Xiong:
Style Projected Clustering for Domain Generalized Semantic Segmentation. 3061-3071
- Shiqi Huang, Tingfa Xu, Ning Shen, Feng Mu, Jianan Li:
Rethinking Few-Shot Medical Segmentation: A Vector Quantization View. 3072-3081
- Lanyun Zhu, Tianrun Chen, Jianxiong Yin, Simon See, Jun Liu:
Continual Semantic Segmentation with Automatic Memory Sample Selection. 3082-3092
- Lixiang Ru, Heliang Zheng, Yibing Zhan, Bo Du:
Token Contrast for Weakly-Supervised Semantic Segmentation. 3093-3102
- Rixin Zhou, Jiafu Wei, Qian Zhang, Ruihua Qi, Xi Yang, Chuntao Li:
Multi-Granularity Archaeological Dating of Chinese Bronze Dings Based on a Knowledge-Guided Relation Graph. 3103-3113
- Xiaoyang Wang , Bingfeng Zhang , Limin Yu, Jimin Xiao:
Hunting Sparsity: Density-Guided Contrastive Learning for Semi-Supervised Semantic Segmentation. 3114-3123
- Xudong Wang, Rohit Girdhar, Stella X. Yu, Ishan Misra:
Cut and Learn for Unsupervised Object Detection and Instance Segmentation. 3124-3134
- Zhaozheng Chen, Qianru Sun:
Extracting Class Activation Maps from Non-Discriminative Features as well. 3135-3144
- Tianheng Cheng, Xinggang Wang, Shaoyu Chen, Qian Zhang, Wenyu Liu:
BoxTeacher: Exploring High-Quality Pseudo Labels for Weakly Supervised Instance Segmentation. 3145-3154
- Xiao Guo, Xiaohong Liu, Zhiyuan Ren, Steven Grosz, Iacopo Masi, Xiaoming Liu:
Hierarchical Fine-Grained Image Forgery Detection and Localization. 3155-3165
- Pei Wang, Nuno Vasconcelos:
Towards Professional Level Crowd Annotation of Expert Domain Data. 3166-3175
- Oriane Siméoni, Chloé Sekkat, Gilles Puy, Antonín Vobecký, Éloi Zablocki, Patrick Pérez:
Unsupervised Object Localization: Observing the Background to Discover Objects. 3176-3186
- Enrico Fini, Pietro Astolfi, Karteek Alahari, Xavier Alameda-Pineda, Julien Mairal, Moin Nabi, Elisa Ricci:
Semi-supervised learning made simple with self-supervised clustering. 3187-3197
- Henri De Plaen, Pierre-François De Plaen, Johan A. K. Suykens , Marc Proesmans, Tinne Tuytelaars , Luc Van Gool:
Unbalanced Optimal Transport: A Unified Framework for Object Detection. 3198-3207
- Jiawei Ma, Yulei Niu, Jincheng Xu, Shiyuan Huang, Guangxing Han, Shih-Fu Chang:
DiGeo: Discriminative Geometry-Aware Learning for Generalized Few-Shot Object Detection. 3208-3218
- Vudit Vudit, Martin Engilberge, Mathieu Salzmann:
CLIP the Gap: A Single Domain Generalization Approach for Object Detection. 3219-3229
- Wenteng Liang, Feng Xue, Yihao Liu, Guofeng Zhong, Anlong Ming:
Unknown Sniffer for Object Detection: Don't Turn a Blind Eye to Unknown Objects. 3230-3239
- Xinjiang Wang, Xingyi Yang , Shilong Zhang, Yijiang Li, Litong Feng, Shijie Fang, Chengqi Lyu, Kai Chen, Wayne Zhang:
Consistent-Teacher: Towards Reducing Inconsistent Pseudo-Targets in Semi-Supervised Object Detection. 3240-3249
- Xiaolin Song, Binghui Chen, Pengyu Li, Jun-Yan He, Biao Wang, Yifeng Geng, Xuansong Xie, Honggang Zhang:
Optimal Proposal Learning for Deployable End-to-End Pedestrian Detection. 3250-3260
- Yipeng Gao , Kun-Yu Lin , Junkai Yan, Yaowei Wang, Wei-Shi Zheng:
AsyFOD: An Asymmetric Adaptation Paradigm for Few-Shot Domain Adaptive Object Detection. 3261-3271
- Chenxi Zheng, Bangzhen Liu, Huaidong Zhang, Xuemiao Xu, Shengfeng He :
Where is My Spot? Few-shot Image Generation via Latent Subspace Optimization. 3272-3281
- Fan Lu, Kai Zhu, Wei Zhai, Kecheng Zheng, Yang Cao:
Uncertainty-Aware Optimal Transport for Semantically Coherent Out-of-Distribution Detection. 3282-3291
- Ronald Xie, Kuan Pang, Gary D. Bader, Bo Wang:
MAESTER: Masked Autoencoder Guided Segmentation at Pixel Resolution for Accurate, Self-Supervised Subcellular Structure Recognition. 3292-3301
- Heng Cai, Shumeng Li, Lei Qi, Qian Yu, Yinghuan Shi, Yang Gao:
Orthogonal Annotation Benefits Barely-supervised Medical Image Segmentation. 3302-3311
- Donghao Zhou, Chunbin Gu, Junde Xu, Furui Liu, Qiong Wang, Guangyong Chen, Pheng-Ann Heng:
RepMode: Learning to Re-Parameterize Diverse Experts for Subcellular Structure Prediction. 3312-3322
- Shahira Abousamra, Rajarsi Gupta, Tahsin M. Kurç, Dimitris Samaras, Joel H. Saltz, Chao Chen:
Topology-Guided Multi-Class Cell Context Generation for Digital Pathology. 3323-3333
- Mingjie Li, Bingqian Lin, Zicong Chen, Haokun Lin, Xiaodan Liang, Xiaojun Chang :
Dynamic Graph Enhanced Contrastive Learning for Chest X-Ray Report Generation. 3334-3343



- Mingu Kang, Heon Song, Seonwook Park , Donggeun Yoo, Sérgio Pereira :
Benchmarking Self-Supervised Learning on Diverse Pathology Datasets. 3344-3354
- Kangning Liu, Weicheng Zhu, Yiqiu Shen, Sheng Liu, Narges Razavian, Krzysztof J. Geras, Carlos Fernandez-Granda:
Multiple Instance Learning via Iterative Self-Paced Supervised Contrastive Learning. 3355-3365
- Rajshekhar Das, Yonatan Dukler, Avinash Ravichandran, Ashwin Swaminathan:
Learning Expressive Prompting With Residuals for Vision Transformers. 3366-3377
- Bartłomiej Olber, Krystian Radlak, Adam Popowicz, Michał Szczepankiewicz, Krystian Chachula:
Detection of Out-of-Distribution Samples Using Binary Neuron Activation Patterns. 3378-3387
- Zihan Zhang, Xiang Xiang:
Decoupling MaxLogit for Out-of-Distribution Detection. 3388-3397
- Zixuan Ding, Ao Wang, Hui Chen, Qiang Zhang, Pengzhang Liu, Yongjun Bao, Weipeng Yan, Jungong Han:
Exploring Structured Semantic Prior for Multi Label Recognition with Incomplete Labels. 3398-3407
- Youngwook Kim , Jae-Myung Kim, Jieun Jeong, Cordelia Schmid, Zeynep Akata, Jungwoo Lee:
Bridging the Gap Between Model Explanations in Partially Annotated Multi-Label Classification. 3408-3417
- Ioannis Maniatis Metaxas, Georgios Tzimiropoulos, Ioannis Patras:
DivClust: Controlling Diversity in Deep Clustering. 3418-3428
- Furen Zhuang, Pierre Moulin:
Deep Semi-Supervised Metric Learning with Mixed Label Propagation. 3429-3438
- Maria Sofia Bucarelli , Lucas Cassano, Federico Siciliano , Amin Mantrach, Fabrizio Silvestri:
Leveraging Inter-Rater Agreement for Classification in the Presence of Noisy Labels. 3439-3448
- Wenbin Li, Zhichen Fan, Jing Huo, Yang Gao:
Modeling Inter-Class and Intra-Class Constraints in Novel Class Discovery. 3449-3458
- Muli Yang, Liancheng Wang, Cheng Deng, Hanwang Zhang:
Bootstrap Your Own Prior: Towards Distribution-Agnostic Novel Class Discovery. 3459-3468
- Tong Wei, Kai Gan:
Towards Realistic Long-Tailed Semi-Supervised Learning: Consistency is All You Need. 3469-3478
- Sheng Zhang, Salman H. Khan, Zhiqiang Shen, Muzammal Naseer, Guangyi Chen, Fahad Shahbaz Khan:
PromptCAL: Contrastive Affinity Learning via Auxiliary Prompts for Generalized Novel Category Discovery. 3479-3488
- Jianqing Xu, Shen Li, Ailin Deng, Miao Xiong, Jiaying Wu, Jiaxiang Wu, Shouhong Ding, Bryan Hooi:
Probabilistic Knowledge Distillation of Face Ensembles. 3489-3498
- Zhipeng Zhou , Lanqing Li, Peilin Zhao, Pheng-Ann Heng, Wei Gong:
Class-Conditional Sharpness-Aware Minimization for Deep Long-Tailed Recognition. 3499-3509
- Yuchen Liu, Yaoming Wang, Yabo Chen, Wenrui Dai, Chenglin Li, Junni Zou, Hongkai Xiong:
Promoting Semantic Connectivity: Dual Nearest Neighbors Contrastive Learning for Unsupervised Domain Generalization. 3510-3519
- Vibashan VS, Poojan Oza, Vishal M. Patel:
Instance Relation Graph Guided Source-Free Domain Adaptive Object Detection. 3520-3530
- You-Wei Luo , Chuan-Xian Ren:
MOT: Masked Optimal Transport for Partial Domain Adaptation. 3531-3540
- Hao Yu, Xu Cheng, Wei Peng :
TOPLight: Lightweight Neural Networks with Task-Oriented Pretraining for Visible-Infrared Recognition. 3541-3550
- Ye Liu, Lingfeng Qiao, Changchong Lu, Di Yin, Chen Lin, Haoyuan Peng, Bo Ren:
OSAN: A One-Stage Alignment Network to Unify Multimodal Alignment and Unsupervised Domain Adaptation. 3551-3560
- Jinjing Zhu, Haotian Bai, Lin Wang :
Patch-Mix Transformer for Unsupervised Domain Adaptation: A Game Perspective. 3561-3571
- Yizhi Wang, Zeyu Huang, Ariel Shamir, Hui Huang, Hao Zhang, Ruizhen Hu:
ARO-Net: Learning Implicit Fields from Anchored Radial Observations. 3572-3581
- Dhanajit Brahma, Piyush Rai:
A Probabilistic Framework for Lifelong Test-Time Adaptation. 3582-3591
- Runpeng Yu, Songhua Liu, Xingyi Yang , Xinchao Wang:
Distribution Shift Inversion for Out-of-Distribution Prediction. 3592-3602
- Jiali Cui, Ying Nian Wu, Tian Han:
Learning Joint Latent Space EBM Prior Model for Multi-layer Generator. 3603-3612
- Saachi Jain, Hadi Salman, Alaa Khaddaj, Eric Wong, Sung Min Park, Aleksander Madry:
A Data-Based Perspective on Transfer Learning. 3613-3622
- Achin Jain, Gurumurthy Swaminathan, Paolo Favaro, Hao Yang, Avinash Ravichandran, Hrayr Harutyunyan, Alessandro Achille, Onkar Dabeer, Bernt Schiele, Ashwin Swaminathan, Stefano Soatto:
A Meta-Learning Approach to Predicting Performance and Data Requirements. 3623-3632



- Hao Li, Charless C. Fowlkes, Hao Yang, Onkar Dabeer, Zhuowen Tu, Stefano Soatto:
Guided Recommendation for Model Fine-Tuning. 3633-3642
- Peng Liao , Yaochu Jin, Wenli Du:
EMT-NAS: Transferring architectural knowledge between tasks from different datasets. 3643-3653
- Runqi Wang, Xiaoyue Duan, Guoliang Kang, Jianzhuang Liu, Shaohui Lin, Songcen Xu, Jinhua Lv, Baochang Zhang:
AttriCLIP: A Non-Incremental Learner for Incremental Knowledge Learning. 3654-3663
- Iordanis Fostiroopoulos, Jiaye Zhu, Laurent Itti:
Batch Model Consolidation: A Multi-Task Model Consolidation Framework. 3664-3676
- Yinglong Wang, Chao Ma, Jianzhuang Liu:
SmartAssign: Learning A Smart Knowledge Assignment Strategy for Deraining and Desnowing. 3677-3686
- Sucheng Ren, Fangyun Wei, Zheng Zhang, Han Hu:
TinyMIM: An Empirical Study of Distilling MIM Pre-trained Models. 3687-3697
- Ameya Prabhu, Hasan Abed Al Kader Hammoud , Puneet K. Dokania, Philip H. S. Torr, Ser-Nam Lim, Bernard Ghanem , Adel Bibi :
Computationally Budgeted Continual Learning: What Does Matter? 3698-3707
- Kangyang Luo, Xiang Li, Yunshi Lan, Ming Gao:
GradMA: A Gradient-Memory-based Accelerated Federated Learning with Alleviated Catastrophic Forgetting. 3708-3717
- Zhen Zhao, Zhizhong Zhang, Xin Tan, Jun Liu, Yanyun Qu, Yuan Xie, Lizhuang Ma:
Rethinking Gradient Projection Continual Learning: Stability/Plasticity Feature Space Decoupling. 3718-3727
- Yushun Tang, Ce Zhang , Heng Xu , Shuoshuo Chen, Jie Cheng, Luziwei Leng, Qinghai Guo, Zhihai He:
Neuro-Modulated Hebbian Learning for Fully Test-Time Adaptation. 3728-3738
- George Cazenavette, Tongzhou Wang, Antonio Torralba, Alexei A. Efros, Jun-Yan Zhu:
Generalizing Dataset Distillation via Deep Generative Prior. 3739-3748
- Jiawei Du, Yidi Jiang, Vincent Y. F. Tan, Joey Tianyi Zhou, Haizhou Li:
Minimizing the Accumulated Trajectory Error to Improve Dataset Distillation. 3749-3758
- Songhua Liu, Jingwen Ye, Rupeng Yu, Xinchao Wang:
Slimmable Dataset Condensation. 3759-3768
- Pengfei Wang, Zhaoxiang Zhang, Zhen Lei, Lei Zhang:
Sharpness-Aware Gradient Matching for Domain Generalization. 3769-3778
- Wonhyeok Choi, Sunghoon Im:
Dynamic Neural Network for Multi-Task Learning Searching across Diverse Network Topologies. 3779-3788
- Ahmed Imtiaz Humayun, Randall Balestriero, Guha Balakrishnan, Richard G. Baraniuk:
SplineCam: Exact Visualization and Characterization of Deep Network Geometry and Decision Boundaries. 3789-3798
- Jaeill Kim, Suhyun Kang, Duhun Hwang, Jungwook Shin, Wonjong Rhee:
VNE: An Effective Method for Improving Deep Representation by Manipulating Eigenvalue Distribution. 3799-3810
- Yuedong Yang, Guihong Li, Radu Marculescu:
Efficient On-Device Training via Gradient Filtering. 3811-3820
- Tang Li, Fengchun Qiao, Mengmeng Ma, Xi Peng:
Are Data-Driven Explanations Robust Against Out-of-Distribution Data? 3821-3831
- Jongin Lim, Youngdong Kim, Byungjai Kim, Chanho Ahn, Jinwoo Shin, Eunho Yang, Seungju Han :
BiasAdv: Bias-Adversarial Augmentation for Model Debiasing. 3832-3841
- Sheng Xu, Yanjing Li, Mingbao Lin, Peng Gao, Guodong Guo, Jinhua Lü, Baochang Zhang:
Q-DETR: An Efficient Low-Bit Quantized Detection Transformer. 3842-3851
- Juncheol Shin, Junhyuk So, Sein Park, Seungyeop Kang, Sungjoo Yoo, Eunhyeok Park:
NIPQ: Noise proxy-based Integrated Pseudo-Quantization. 3852-3861
- Vinu Sankar Sadasivan, Mahdi Soltanolkotabi , Soheil Feizi:
CUDA: Convolution-Based Unlearnable Datasets. 3862-3871
- Kaiwen Cui, Yingchen Yu, Fangneng Zhan, Shengcui Liao, Shijian Lu, Eric P. Xing:
KD-DLGAN: Data Limited Image Generation via Knowledge Distillation. 3872-3882
- Siddarth Asokan , Chandra Sekhar Seelamantula:
Spider GAN: Leveraging Friendly Neighbors to Accelerate GAN Training. 3883-3893
- Harleen Hanspal, Alessio Lomuscio:
Efficient Verification of Neural Networks Against LVM-Based Specifications. 3894-3903
- Kexin Sun, Zhineng Chen, Gongwei Wang, Jun Liu, Xiongjun Ye, Yu-Gang Jiang:
Bi-directional Feature Fusion Generative Adversarial Network for Ultra-high Resolution Pathological Image Virtual Re-staining. 3904-3913
- Xuan Zhang, Shiyu Li, Xi Li, Ping Huang, Jilong Shan, Ting Chen:
DeSTSeg: Segmentation Guided Denoising Student-Teacher for Anomaly Detection. 3914-3923



- Ying Zhao : **OmniAL: A Unified CNN Framework for Unsupervised Anomaly Localization.** 3924-3933
- Jiahua Dong, Duzhen Zhang, Yang Cong, Wei Cong, Henghui Ding , Dengxin Dai: **Federated Incremental Semantic Segmentation.** 3934-3943
- Sangmook Kim, Sangmin Bae, Hwanjun Song, Se-Young Yun: **Re-Thinking Federated Active Learning Based on Inter-Class Diversity.** 3944-3953
- Ruipeng Zhang, Qinwei Xu, Jiangchao Yao, Ya Zhang , Qi Tian, Yanfeng Wang: **Federated Domain Generalization with Generalization Adjustment.** 3954-3963
- Bo Li, Mikkel N. Schmidt , Tommy S. Alstrøm , Sebastian U. Stich: **On the Effectiveness of Partial Variance Reduction in Federated Learning with Heterogeneous Data.** 3964-3973
- Joshua C. Zhao, Ahmed Roushdy Elkordy, Atul Sharma, Yahya H. Ezzeldin, Salman Avestimehr, Saurabh Bagchi: **The Resource Problem of Using Linear Layer Leakage Attack in Federated Learning.** 3974-3983
- Jiaming Zhang, Xingjun Ma, Qi Yi, Jitao Sang, Yu-Gang Jiang, Yaowei Wang, Changsheng Xu: **Unlearnable Clusters: Towards Label-Agnostic Unlearnable Examples.** 3984-3993
- Shichao Dong, Jin Wang, Renhe Ji, Jiajun Liang, Haoqiang Fan, Zheng Ge: **Implicit Identity Leakage: The Stumbling Block to Improving Deepfake Detection Generalization.** 3994-4004
- Kuofeng Gao, Yang Bai , Jindong Gu, Yong Yang, Shu-Tao Xia: **Backdoor Defense via Adaptively Splitting Poisoned Dataset.** 4005-4014
- Sheng-Yen Chou , Pin-Yu Chen, Tsung-Yi Ho: **How to Backdoor Diffusion Models?** 4015-4024
- Mengxin Zheng, Qian Lou, Lei Jiang: **TrojViT: Trojan Insertion in Vision Transformers.** 4025-4034
- Weixin Chen, Dawn Song, Bo Li: **TrojDiff: Trojan Attacks on Diffusion Models with Diverse Targets.** 4035-4044
- Zikui Cai, Yaoteng Tan, M. Salman Asif: **Ensemble-based Blackbox Attacks on Dense Prediction.** 4045-4055
- Yunrui Yu, Cheng-Zhong Xu: **Efficient Loss Function by Minimizing the Detrimental Effect of Floating-Point Errors on Gradient-Based Attacks.** 4056-4066
- Iuri Frosio, Jan Kautz: **The Best Defense is a Good Offense: Adversarial Augmentation Against Adversarial Attacks.** 4067-4076
- Minjing Dong , Chang Xu: **Adversarial Robustness via Random Projection Filters.** 4077-4086
- Bilel Tarchoun, Anouar Ben Khalifa, Mohamed Ali Mahjoub, Nael B. Abu-Ghazaleh, Ihsen Alouani : **Jedi: Entropy-Based Localization and Removal of Adversarial Patches.** 4087-4095
- Aishan Liu, Shiyu Tang, Siyuan Liang, Ruihao Gong, Boxi Wu, Xianglong Liu, Dacheng Tao: **Exploring the Relationship Between Architectural Design and Adversarially Robust Generalization.** 4096-4107
- Yong Guo, David Stutz, Bernt Schiele: **Improving Robustness of Vision Transformers by Reducing Sensitivity to Patch Corruptions.** 4108-4118
- Xiao Yang, Chang Liu, Longlong Xu , Yikai Wang, Yinpeng Dong, Ning Chen, Hang Su, Jun Zhu: **Towards Effective Adversarial Textured 3D Meshes on Physical Face Recognition.** 4119-4128
- Zhendong Wang, Jianmin Bao, Wengang Zhou, Weilun Wang, Houqiang Li: **AltFreezing for More General Video Face Forgery Detection.** 4129-4138
- Alankar Kotwal, Anat Levin, Ioannis Gkioulekas : **Passive Micron-Scale Time-of-Flight with Sunlight Interferometry.** 4139-4149
- Peng Wang, Yuan Liu, Zhaoxi Chen, Lingjie Liu, Ziwei Liu, Taku Komura, Christian Theobalt, Wenping Wang: **F²-NeRF: Fast Neural Radiance Field Training with Free Camera Trajectories.** 4150-4159
- Wenjing Bian, Zirui Wang, Kejie Li, Jia-Wang Bian: **NoPe-NeRF: Optimising Neural Radiance Field with No Pose Prior.** 4160-4169
- Peng Wang , Lingzhe Zhao , Ruijie Ma, Peidong Liu: **BAD-NeRF: Bundle Adjusted Deblur Neural Radiance Fields.** 4170-4179
- Jamie Wynn, Daniyar Turmukhambetov: **DiffusioNeRF: Regularizing Neural Radiance Fields with Denoising Diffusion Models.** 4180-4189
- Prune Truong, Marie-Julie Rakotosaona, Fabian Manhardt, Federico Tombari: **SPARF: Neural Radiance Fields from Sparse and Noisy Poses.** 4190-4200
- Rahul Goel, Dhawal Sirikonda, Saurabh Saini, P. J. Narayanan: **Interactive Segmentation of Radiance Fields.** 4201-4211
- Sungheon Park, Minjung Son , Seokhwan Jang, Young Chun Ahn, Ji-Yeon Kim, Nahyup Kang: **Temporal Interpolation is all You Need for Dynamic Neural Radiance Fields.** 4212-4221



- Lingzhi Li, Zhen Shen, Zhongshu Wang, Li Shen, Liefeng Bo:
Compressing Volumetric Radiance Fields to 1 MB. 4222-4231
- Kang Han, Wei Xiang:
Multiscale Tensor Decomposition and Rendering Equation Encoding for View Synthesis. 4232-4241
- Yuechen Zhang, Zexin He, Jinbo Xing, Xufeng Yao, Jiaya Jia:
Ref-NPR: Reference-Based Non-Photorealistic Radiance Fields for Controllable Scene Stylization. 4242-4251
- Sida Peng, Yunzhi Yan, Qing Shuai, Hujun Bao, Xiaowei Zhou:
Representing Volumetric Videos as Dynamic MLP Maps. 4252-4262
- Wei Dong, Christopher B. Choy, Charles Loop, Or Litany, Yuke Zhu, Anima Anandkumar:
Fast Monocular Scene Reconstruction with Global-Sparse Local-Dense Grids. 4263-4272
- Zhengqi Li, Qianqian Wang, Forrester Cole, Richard Tucker, Noah Snavely:
DynIBaR: Neural Dynamic Image-Based Rendering. 4273-4284
- Michael Fischer, Tobias Ritschel:
Plateau-Reduced Differentiable Path Tracing. 4285-4294
- Haoqian Wu, Zhipeng Hu, Lincheng Li, Yongqiang Zhang , Changjie Fan, Xin Yu :
NeFII: Inverse Rendering for Reflectance Decomposition with Near-Field Indirect Illumination. 4295-4304
- Ziang Cheng, Junxuan Li, Hongdong Li:
WildLight: In-the-wild Inverse Rendering with a Flashlight. 4305-4314
- Taotao Zhou, Kai He, Di Wu, Teng Xu, Qixuan Zhang, Kuixiang Shao, Wenzheng Chen, Lan Xu, Jingyi Yu:
Relightable Neural Human Assets from Multi-view Gradient Illuminations. 4315-4327
- Norman Müller, Yawar Siddiqui, Lorenzo Porzi, Samuel Rota Bulò, Peter Kontschieder, Matthias Nießner:
DiffRF: Rendering-Guided 3D Radiance Field Diffusion. 4328-4338
- Tianyuan Zhang, Mark Sheinin, Dorian Chan, Mark Rau, Matthew O'Toole, Srinivasa G. Narasimhan :
Analyzing Physical Impacts Using Transient Surface Wave Imaging. 4339-4348
- Byeongjoo Ahn, Michael DeZeeuw, Ioannis Gkioulekas , Aswin C. Sankaranarayanan:
Neural Kaleidoscopic Space Sculpting. 4349-4358
- Yongqiang Zhang , Zhipeng Hu, Haoqian Wu, Minda Zhao, Lincheng Li, Zhengxia Zou, Changjie Fan:
Towards Unbiased Volume Rendering of Neural Implicit Surfaces with Geometry Priors. 4359-4368
- Jiahui Huang, Zan Gojcic, Matan Atzmon, Or Litany, Sanja Fidler, Francis Williams:
Neural Kernel Surface Reconstruction. 4369-4379
- Mingye Xu, Mutian Xu, Tong He, Wanli Ouyang, Yali Wang, Xiaoguang Han, Yu Qiao:
MM-3DSce: 3D Scene Understanding by Customizing Masked Modeling with Informative-Preserved Reconstruction and Self-Distilled Consistency. 4380-4390
- Dario Pavllo, David Joseph Tan, Marie-Julie Rakotosaona, Federico Tombari:
Shape, Pose, and Appearance from a Single Image via Bootstrapped Radiance Field Inversion. 4391-4401
- Yinghao Xu, Menglei Chai, Zifan Shi, Sida Peng, Ivan Skorokhodov , Aliaksandr Siarohin, Ceyuan Yang, Yujun Shen, Hsin-Ying Lee, Bolei Zhou, Sergey Tulyakov:
DisCoScene: Spatially Disentangled Generative Radiance Fields for Controllable 3D-aware Scene Synthesis. 4402-4412
- Chi-Chong Wong:
Heat Diffusion Based Multi-Scale and Geometric Structure-Aware Transformer for Mesh Segmentation. 4413-4422
- Yu Deng, Baoyuan Wang, Heung-Yeung Shum:
Learning Detailed Radiance Manifolds for High-Fidelity and 3D-Consistent Portrait Synthesis from Monocular Image. 4423-4433
- Kangle Deng , Gengshan Yang, Deva Ramanan, Jun-Yan Zhu:
3D-aware Conditional Image Synthesis. 4434-4445
- Anna Frühstück, Nikolaos Sarafianos, Yuanlu Xu, Peter Wonka, Tony Tung:
VIVE3D: Viewpoint-Independent Video Editing using 3D-Aware GANs. 4446-4455
- Yen-Chi Cheng, Hsin-Ying Lee, Sergey Tulyakov, Alexander G. Schwing, Liangyan Gui:
SDFusion: Multimodal 3D Shape Completion, Reconstruction, and Generation. 4456-4465
- Konstantinos Tertikas, Despoina Paschalidou , Boxiao Pan, Jeong Joon Park, Mikaela Angelina Uy, Ioannis Z. Emiris, Yannis Avrithis, Leonidas J. Guibas:
Generating Part-Aware Editable 3D Shapes without 3D Supervision. 4466-4478
- Dejia Xu, Yifan Jiang, Peihao Wang, Zhiwen Fan, Yi Wang, Zhangyang Wang:
NeuralLift-360: Lifting an in-the-Wild 2D Photo to A 3D Object with 360° Views. 4479-4489
- Baojin Huang, Zhongyuan Wang, Jifan Yang , Jiaxin Ai, Qin Zou, Qian Wang, Dengpan Ye:
Implicit Identity Driven Deepfake Face Swapping Detection. 4490-4499
- Rohith Agaram, Shaurya Dewan, Rahul Sajnani, Adrien Poulenard, K. Madhava Krishna, Srinath Sridhar:
Canonical Fields: Self-Supervised Learning of Pose-Canonicalized Neural Fields. 4500-4510
- Xingyu Ren , Jiankang Deng, Chao Ma, Yichao Yan, Xiaokang Yang:
Improving Fairness in Facial Albedo Estimation via Visual-Textual Cues. 4511-4520



- Menghua Wu, Hao Zhu, Linjia Huang, Yiyu Zhuang, Yuanxun Lu, Xun Cao:
High-fidelity 3D Face Generation from Natural Language Descriptions. 4521-4530
- Heyuan Li, Bo Wang , Yu Cheng, Mohan S. Kankanhalli , Robby T. Tan:
DSFNet: Dual Space Fusion Network for Occlusion-Robust 3D Dense Face Alignment. 4531-4540
- Yunpeng Bai, Yanbo Fan, Xuan Wang, Yong Zhang, Jingxiang Sun, Chun Yuan, Ying Shan:
High-fidelity Facial Avatar Reconstruction from Monocular Video with Generative Priors. 4541-4551
- Rameen Abdal, Hsin-Ying Lee, Peihao Zhu, Menglei Chai, Aliaksandr Siarohin, Peter Wonka, Sergey Tulyakov:
3DAvatarGAN: Bridging Domains for Personalized Editable Avatars. 4552-4562
- Tengfei Wang, Bo Zhang, Ting Zhang, Shuyang Gu, Jianmin Bao, Tadas Baltrusaitis, Jingjing Shen, Dong Chen, Fang Wen, Qifeng Chen, Baining Guo:
RODIN: A Generative Model for Sculpting 3D Digital Avatars Using Diffusion. 4563-4573
- Wojciech Zielezny, Timo Bolkart, Justus Thies:
Instant Volumetric Head Avatars. 4574-4584
- Siddarth Ravichandran, Ondrej Texler, Dimitar Dinev, Hyun Jae Kang:
Synthesizing Photorealistic Virtual Humans Through Cross-Modal Disentanglement. 4585-4594
- Xingyi Li , Zhiguo Cao, Huiqiang Sun, Jianming Zhang, Ke Xian, Guosheng Lin:
3D Cinematography from a Single Image. 4595-4605
- Luyang Zhu, Dawei Yang, Tyler Zhu, Fitsum Reda, William Chan, Chitwan Saharia, Mohammad Norouzi, Ira Kemelmacher-Shlizerman:
TryOnDiffusion: A Tale of Two UNets. 4606-4615
- Xingqun Qi, Chen Liu, Muyi Sun, Lincheng Li, Changjie Fan, Xin Yu :
Diverse 3D Hand Gesture Prediction from Body Dynamics by Bilateral Hand Disentanglement. 4616-4626
- Yasamin Jafarian, Tuanfeng Y. Wang, Duygu Ceylan, Jimei Yang, Nathan Carr, Yi Zhou, Hyun Soo Park:
Normal-guided Garment UV Prediction for Human Re-texturing. 4627-4636
- Lingteng Qiu, Guanying Chen , Jiapeng Zhou, Mutian Xu, Junle Wang, Xiaoguang Han:
REC-MV: REconstructing 3D Dynamic Cloth from Monocular Videos. 4637-4646
- Yukang Cao, Kai Han, Kwan-Yee K. Wong:
SeSDF: Self-Evolved Signed Distance Field for Implicit 3D Clothed Human Reconstruction. 4647-4657
- Rolando Alexopoulos Potamias, Stylianos Ploumpis, Stylianos Moschoglou, Vasileios Triantafyllou, Stefanos Zafeiriou:
Handy: Towards a High Fidelity 3D Hand Shape and Appearance Model. 4670-4680
- Nikolas Lamb, Cameron Palmer, Benjamin Molloy, Sean Banerjee, Natasha Kholgade Banerjee:
Fantastic Breaks: A Dataset of Paired 3D Scans of Real-World Broken Objects and Their Complete Counterparts. 4681-4691
- Jeff Tan, Gengshan Yang, Deva Ramanan:
Distilling Neural Fields for Real-Time Articulated Shape Reconstruction. 4692-4701
- Rui Guo, Jasmine Collins, Oscar de Lima, Andrew Owens:
GANmouflage: 3D Object Nondetection with Texture Fields. 4702-4712
- Shashank Tripathi, Lea Müller, Chun-Hao P. Huang, Omid Taheri, Michael J. Black, Dimitrios Tzionas :
3D Human Pose Estimation via Intuitive Physics. 4713-4725
- Ilya A. Petrov , Riccardo Marin, Julian Chibane, Gerard Pons-Moll:
Object pop-up: Can we infer 3D objects and their poses from human interactions alone? 4726-4736
- Yinzhen Xu, Weikang Wan, Jialiang Zhang, Haoran Liu, Zikang Shan, Hao Shen, Ruicheng Wang, Haoran Geng, Yijia Weng, Jiayi Chen, Tengyu Liu, Li Yi, He Wang:
UniDexGrasp: Universal Robotic Dexterous Grasping via Learning Diverse Proposal Generation and Goal-Conditioned Policy. 4737-4746
- Xiongbiao Luo:
Constrained Evolutionary Diffusion Filter for Monocular Endoscope Tracking. 4747-4756
- Xianghui Xie, Bharat Lal Bhatnagar, Gerard Pons-Moll:
Visibility Aware Human-Object Interaction Tracking from Single RGB Camera. 4757-4768
- Hoseong Cho, Chanwoo Kim, Jihyeon Kim, Seongyeong Lee, Elkhan Ismayilzada , Seungryul Baek:
Transformer-based Unified Recognition of Two Hands Manipulating Objects. 4769-4778
- Akash Sengupta, Ignas Budvytis, Roberto Cipolla:
HuManiFlow: Ancestor-Conditioned Normalising Flows on SO(3) Manifolds for Human Pose and Shape Distribution Estimation. 4779-4789
- Zhenhua Tang, Zhaofan Qiu, Yanbin Hao, Richang Hong, Ting Yao:
3D Human Pose Estimation with Spatio-Temporal Criss-Cross Attention. 4790-4799
- Hai Ci, Mingdong Wu, Wentao Zhu, Xiaoxuan Ma, Hao Dong, Fangwei Zhong, Yizhou Wang:
GFPose: Learning 3D Human Pose Prior with Gradient Fields. 4800-4810
- Edward Vendrow, Duy-Tho Le, Jianfei Cai, Hamid Rezatofighi:
JRDB-Pose: A Large-Scale Dataset for Multi-Person Pose Estimation and Tracking. 4811-4820



- Qiyuan He, Linlin Yang, Kerui Gu, Qiuxia Lin, Angela Yao:
Analyzing and Diagnosing Pose Estimation with Attributions. 4821-4830
- Yang Hai, Rui Song, Jiaoqiao Li, Yinlin Hu:
Shape-Constraint Recurrent Flow for 6D Object Pose Estimation. 4831-4840
- Hanzhi Chen, Fabian Manhardt, Nassir Navab, Benjamin Busam:
TexPose: Neural Texture Learning for Self-Supervised 6D Object Pose Estimation. 4841-4852
- Chun-Han Yao, Wei-Chih Hung, Yuanzhen Li, Michael Rubinstein, Ming-Hsuan Yang, Varun Jampani:
Hi-LASSIE: High-Fidelity Articulated Shape and Skeleton Discovery from Sparse Image Ensemble. 4853-4862
- Bangyan Liao, Delin Qu, Yifei Xue, Huiqing Zhang, Yizhen Lao:
Revisiting Rolling Shutter Bundle Adjustment: Toward Accurate and Fast Solution. 4863-4871
- Yaqing Ding, Jian Yang, Viktor Larsson, Carl Olsson, Kalle Åström :
Revisiting the P3P Problem. 4872-4880
- Samarth Sinha, Roman Shapovalov, Jeremy Reizenstein, Ignacio Rocco, Natalia Neverova, Andrea Vedaldi, David Novotný:
Common Pets in 3D: Dynamic New-View Synthesis of Real-Life Deformable Categories. 4881-4891
- Kejie Li, Jia-Wang Bian, Robert Castle, Philip H. S. Torr, Victor Adrian Prisacariu:
MobileBrick: Building LEGO for 3D Reconstruction on Mobile Devices. 4892-4901
- Jiahui Lei, Congyue Deng, Karl Schmeckpeper, Leonidas J. Guibas, Kostas Daniilidis:
EFEM: Equivariant Neural Field Expectation Maximization for 3D Object Segmentation Without Scene Supervision. 4902-4912
- Bokui Shen, Xinchen Yan, Charles R. Qi, Mahyar Najibi, Boyang Deng, Leonidas J. Guibas, Yin Zhou, Dragomir Anguelov:
GINA-3D: Learning to Generate Implicit Neural Assets in the Wild. 4913-4926
- Karmesh Yadav, Ram Ramrakhya, Santhosh Kumar Ramakrishnan, Théophile Gervet, John M. Turner, Aaron Gokaslan, Noah Maestre, Angel Xuan Chang, Dhruv Batra, Manolis Savva, Alexander William Clegg, Devendra Singh Chaplot:
Habitat-Matterport 3D Semantics Dataset. 4927-4936
- Tao Chu, Pan Zhang, Qiong Liu, Jiaqi Wang:
BUOL: A Bottom-Up Framework with Occupancy-Aware Lifting for Panoptic 3D Scene Reconstruction From a Single Image. 4937-4946
- Xinhua Cheng, Yanmin Wu , Mengxi Jia, Qian Wang, Jian Zhang:
Panoptic Compositional Feature Field for Editable Scene Rendering with Network-Inferred Labels via Metric Learning. 4947-4957
- Yash Bhalgat, João F. Henriques, Andrew Zisserman:
A Light Touch Approach to Teaching Transformers Multi-view Geometry. 4958-4969
- Yilun Du, Cameron Smith, Ayush Tewari, Vincent Sitzmann:
Learning to Render Novel Views from Wide-Baseline Stereo Pairs. 4970-4980
- Lukas Mehl, Jenny Schmalfuss, Azin Jahedi, Yaroslava Nalivayko, Andrés Bruhn :
Spring: A High-Resolution High-Detail Dataset and Benchmark for Scene Flow, Optical Flow and Stereo. 4981-4991
- Viktor Rudnev, Mohamed Elgharib, Christian Theobalt, Vladislav Golyanik:
EventNeRF: Neural Radiance Fields from a Single Colour Event Camera. 4992-5002
- Shengjie Zhu , Xiaoming Liu:
LightedDepth: Video Depth Estimation in Light of Limited Inference View Angles. 5003-5012
- Ruicheng Feng, Chongyi Li, Huaijin G. Chen, Shuai Li, Jinwei Gu, Chen Change Loy:
Generating Aligned Pseudo-Supervision from Non-Aligned Data for Image Restoration in Under-Display Camera. 5013-5022
- Donggun Kim, Hyeonjoong Jang, Inchul Kim, Min H. Kim:
Spatio-Focal Bidirectional Disparity Estimation from a Dual-Pixel Image. 5023-5032
- Chao Ning, Hongping Gan:
Trap Attention: Monocular Depth Estimation with Manual Traps. 5033-5043
- Eric Brachmann, Tommaso Cavallari, Victor Adrian Prisacariu:
Accelerated Coordinate Encoding: Learning to Relocalize in Minutes Using RGB and Poses. 5044-5053
- Brevin Tilmon, Zhanghao Sun, Sanjeev J. Koppal, Yicheng Wu, Georgios Evangelidis, Ramzi Zahreddine, Gurunandan Krishnan, Sizhuo Ma, Jian Wang:
Energy-Efficient Adaptive 3D Sensing. 5054-5063
- Shun-Cheng Wu, Keisuke Tateno, Nassir Navab, Federico Tombari:
Incremental 3D Semantic Scene Graph Prediction from RGB Sequences. 5064-5074
- Zhanghao Sun, Wei Ye, Jinhui Xiong, Gyeongmin Choe, Jialiang Wang, Shuochen Su, Rakesh Ranjan:
Consistent Direct Time-of-Flight Video Depth Super-Resolution. 5075-5085
- Chittesh Thavamani, Mengtian Li, Francesco Ferroni, Deva Ramanan:
Learning to Zoom and Unzoom. 5086-5095
- Yuqi Wang, Yuntao Chen, Zhaoxiang Zhang:
FrustumFormer: Adaptive Instance-aware Resampling for Multi-view 3D Detection. 5096-5105



- Jiawei He Yuntao Chen, Naiyan Wang, Zhaoxiang Zhang:
3D Video Object Detection with Learnable Object-Centric Global Optimization. 5106-5115
- Shengchao Zhou, Weizhou Liu, Chen Hu, Shuchang Zhou, Chao Ma:
UniDistill: A Universal Cross-Modality Knowledge Distillation Framework for 3D Object Detection in Bird's-Eye View. 5116-5125
- Haojie Zhao, Junsong Chen, Lijun Wang, Huchuan Lu:
ARKitTrack: A New Diverse Dataset for Tracking Using Mobile RGB-D Data. 5126-5135
- Qi Ming, Lingjuan Miao, Zhe Ma, Lin Zhao, Zhiqiang Zhou, Xuhui Huang, Yuanpei Chen, Yufei Guo:
Deep Dive into Gradients: Better Optimization for 3D Object Detection with Gradient-Corrected IoU Supervision. 5136-5145
- Han Liu Yuhao Wu, Zhiyuan Yu , Yevgeniy Vorobeychik, Ning Zhang:
SlowLiDAR: Increasing the Latency of LiDAR-Based Detection Using Adversarial Examples. 5146-5155
- Nishant Kumar Sinisa Segvic, Abouzar Eslami, Stefan Gumhold:
Normalizing Flow based Feature Synthesis for Outlier-Aware Object Detection. 5156-5165
- Chao Zhou Yanan Zhang , Jiaxin Chen, Di Huang:
OcTr: Octree-Based Transformer for 3D Object Detection. 5166-5175
- Sijie Wang, Qiyu Kang, Rui She , Wei Wang, Kai Zhao, Yang Song, Wee Peng Tay:
HypLiLoc: Towards Effective LiDAR Pose Regression with Hyperbolic Fusion. 5176-5185
- Song Wang, Wentong Li, Wenyu Liu, Xiaolu Liu, Jianke Zhu:
LiDAR2Map: In Defense of LiDAR-Based Semantic Map Construction Using Online Camera Distillation. 5186-5195
- Chenhang He, Ruihuang Li, Yabin Zhang, Shuai Li, Lei Zhang:
MSF: Motion-guided Sequential Fusion for Efficient 3D Object Detection from Point Cloud Sequences. 5196-5205
- Fei Xue, Ignas Budvytis, Roberto Cipolla:
SFD2: Semantic-Guided Feature Detection and Description. 5206-5216
- Lucas Nunes, Louis Wiesmann, Rodrigo Marcuzzi, Xieyuanli Chen , Jens Behley, Cyrill Stachniss:
Temporal Consistent 3D LiDAR Representation Learning for Semantic Perception in Autonomous Driving. 5217-5228
- Bo Pang, Hongchi Xia, Cewu Lu:
Unsupervised 3D Point Cloud Representation Learning by Triangle Constrained Contrast for Autonomous Driving. 5229-5239
- Angelika Ando, Spyros Gidaris, Andrei Bursuc, Gilles Puy, Alexandre Boulch, Renaud Marlet:
RangeViT: Towards Vision Transformers for 3D Semantic Segmentation in Autonomous Driving. 5240-5250
- Yanhao Wu , Tong Zhang, Wei Ke, Sabine Süsstrunk, Mathieu Salzmann:
Spatiotemporal Self-Supervised Learning for Point Clouds in the Wild. 5251-5260
- Utkarsh Mall, Bharath Hariharan, Kavita Bala :
Change-Aware Sampling and Contrastive Learning for Satellite Images. 5261-5270
- Yaqi Shen, Le Hui, Jin Xie, Jian Yang:
Self-Supervised 3D Scene Flow Estimation Guided by Superpoints. 5271-5280
- Itai Lang, Dror Aiger, Forrester Cole, Shai Avidan, Michael Rubinstein:
SCOOP: Self-Supervised Correspondence and Optimization-Based Scene Flow. 5281-5290
- Anthony Chen, Kevin Zhang, Renrui Zhang, Zihan Wang, Yuheng Lu, Yandong Guo, Shanghang Zhang:
PiMAE: Point Cloud and Image Interactive Masked Autoencoders for 3D Object Detection. 5291-5301
- Yaomin Huang, Ning Liu, Zhengping Che, Zhiyuan Xu, Chaomin Shen, Yixin Peng, Guixu Zhang, Xinmei Liu, Feifei Feng, Jian Tang:
CP³: Channel Pruning Plug-in for Point-Based Networks. 5302-5312
- Xiuwei Xu, Ziwei Wang, Jie Zhou, Jiwen Lu:
Binarizing Sparse Convolutional Networks for Efficient Point Cloud Analysis. 5313-5322
- Junming Zhang, Haomeng Zhang, Ram Vasudevan, Matthew Johnson-Roberson:
Hyperspherical Embedding for Point Cloud Completion. 5323-5332
- Chengzhi Wu, Junwei Zheng, Julius Pfrommer, Jürgen Beyerer:
Attention-Based Point Cloud Edge Sampling. 5333-5343
- Renrui Zhang, Liuhui Wang, Yali Wang, Peng Gao, Hongsheng Li , Jianbo Shi:
Starting from Non-Parametric Networks for 3D Point Cloud Analysis. 5344-5353
- Yun He, Danhang Tang, Yinda Zhang, Xiangyang Xue, Yanwei Fu:
Grad-PU: Arbitrary-Scale Point Cloud Upsampling via Gradient Descent with Learned Distance Functions. 5354-5363
- Jiacheng Deng, Chuxin Wang, Jiahao Lu, Jianfeng He, Tianzhu Zhang, Jiyang Yu, Zhe Zhang:
SE-ORNet: Self-Ensembling Orientation-Aware Network for Unsupervised Point Cloud Shape Correspondence. 5364-5373
- Shengwei Qin , Zhong Li, Ligang Liu:
Robust 3D Shape Classification via Non-local Graph Attention Network. 5374-5383



- Hao Yu, Zheng Qin, Ji Hou, Mahdi Saleh, Dongsheng Li, Benjamin Busam, Slobodan Ilic:
Rotation-Invariant Transformer for Point Cloud Matching. 5384-5393
- Zheng Qin, Hao Yu, Changjian Wang, Yuxing Peng, Kai Xu:
Deep Graph-based Spatial Consistency for Robust Non-rigid Point Cloud Registration. 5394-5403
- Tianlu Zhang, Hongyuan Guo, Qiang Jiao, Qiang Zhang, Jungong Han:
Efficient RGB-T Tracking via Cross-Modality Distillation. 5404-5413
- Daniel Barath, Denys Rozumnyi , Ivan Eichhardt, Levente Hajder, Jiri Matas:
Finding Geometric Models by Clustering in the Consensus Space. 5414-5424
- Dihe Huang, Ying Chen, Yong Liu, Jianlin Liu, Shang Xu, Wenlong Wu, Yikang Ding, Fan Tang, Chengjie Wang:
Adaptive Assignment for Geometry Aware Local Feature Matching. 5425-5434
- Zhibo Rao, Bangshu Xiong, Mingyi He , Yuchao Dai, Renjie He, Zhelun Shen, Xing Li:
Masked Representation Learning for Domain Generalized Stereo Matching. 5435-5444
- Han Ling, Yinghui Sun, Quansen Sun, Zhenwen Ren:
Learning Optical Expansion from Scale Matching. 5445-5454
- Hyunyoung Jung , Zhuo Hui, Lei Luo, Haitao Yang, Feng Liu, Sungjoo Yoo, Rakesh Ranjan, Denis Demandolx:
AnyFlow: Arbitrary Scale Optical Flow with Implicit Neural Representation. 5455-5465
- Mohammad Amin Shabani, Sepidehsadat Hosseini, Yasutaka Furukawa:
HouseDiffusion: Vector Floorplan Generation via a Diffusion Model with Discrete and Continuous Denoising. 5466-5475
- Abdul Hannan Khan, Mohammed Shariq Nawaz, Andreas Dengel:
Localized Semantic Feature Mixers for Efficient Pedestrian Detection in Autonomous Driving. 5476-5485
- Haibao Yu, Wenxian Yang, Hongzhi Ruan, Zhenwei Yang, Yingjuan Tang, Xu Gao, Xin Hao, Yifeng Shi, Yifeng Pan, Ning Sun, Juan Song, Jirui Yuan, Ping Luo, Zaiqing Nie:
V2X-Seq: A Large-Scale Sequential Dataset for Vehicle-Infrastructure Cooperative Perception and Forecasting. 5486-5495
- Junru Gu, Chenxu Hu, Tianyuan Zhang, Xuanyao Chen, Yilun Wang, Yue Wang, Hang Zhao:
ViP3D: End-to-End Visual Trajectory Prediction via 3D Agent Queries. 5496-5506
- Dekai Zhu, Guangyao Zhai, Yan Di, Fabian Manhardt, Hendrik Berkemeyer, Tuan Tran, Nassir Navab, Federico Tombari, Benjamin Busam:
IPCC-TP: Utilizing Incremental Pearson Correlation Coefficient for Joint Multi-Agent Trajectory Prediction. 5507-5516
- Weibo Mao, Chenxin Xu, Qi Zhu, Siheng Chen, Yanfeng Wang:
Leapfrog Diffusion Model for Stochastic Trajectory Prediction. 5517-5526
- Xiaoning Sun, Huaijiang Sun, Bin Li, Dong Wei, Weiqing Li, Jianfeng Lu:
DeFeeNet: Consecutive 3D Human Motion Prediction with Deviation Feedback. 5527-5536
- Zhehan Kan, Shuoshuo Chen, Ce Zhang , Yushun Tang, Zhihai He:
Self-Correctable and Adaptable Inference for Generalizable Human Pose Estimation. 5537-5546
- Shiwei Jin, Zhen Wang, Lei Wang, Ning Bi, Truong Q. Nguyen:
ReDirTrans: Latent-to-Latent Translation for Gaze and Head Redirection. 5547-5556
- Zhou Huang, Hang Dai, Tian-Zhu Xiang, Shuo Wang, Huai-Xin Chen, Jie Qin, Huan Xiong:
Feature Shrinkage Pyramid for Camouflaged Object Detection with Transformers. 5557-5566
- Siyuan Li, Tobias Fischer, Lei Ke, Henghui Ding , Martin Danelljan, Fisher Yu:
OVTrack: Open-Vocabulary Multiple Object Tracking. 5567-5577
- Huanzhang Dou, Pengyi Zhang, Wei Su, Yunlong Yu, Yining Lin, Xi Li:
GaitGCI: Generative Counterfactual Intervention for Gait Recognition. 5578-5588
- Dimitrios Kollias:
Multi-Label Compound Expression Recognition: C-EXPR Database & Network. 5589-5598
- Lianxin Xie, Wen Xue, Zhen Xu, Si Wu, Zhiwen Yu, Hau-San Wong:
Blemish-aware and Progressive Face Retouching with Limited Paired Data. 5599-5608
- Yue Gao, Yuan Zhou, Jinglu Wang , Xiao Li, Xiang Ming, Yan Lu:
High-Fidelity and Freely Controllable Talking Head Video Generation. 5609-5619
- Lei Wang , Piotr Koniusz:
3Mformer: Multi-order Multi-mode Transformer for Skeletal Action Recognition. 5620-5631
- Zixiang Zhou, Baoyuan Wang:
UDE: A Unified Driving Engine for Human Motion Generation. 5632-5641
- Nico Messikommer , Carter Fang, Mathias Gehrig, Davide Scaramuzza:
Data-Driven Feature Tracking for Event Cameras. 5642-5651
- Xiaoqian Shen, Xiang Li, Mohamed Elhoseiny:
MoStGAN-V: Video Generation with Temporal Motion Styles. 5652-5661
- Boyang Zhang, Kehua Ma, Suping Wu, Zhixiang Yuan:
Two-stage Co-segmentation Network Based on Discriminative Representation for Recovering Human Mesh from

**Videos.** 5662-5670

- Bin Fan, Yuxin Mao, Yuchao Dai, Zhexiong Wan, Qi Liu : **Joint Appearance and Motion Learning for Efficient Rolling Shutter Correction.** 5671-5681
- Guozhen Zhang, Yuhua Zhu, Haonan Wang, Youxin Chen, Gangshan Wu, Limin Wang: **Extracting Motion and Appearance via Inter-Frame Attention for Efficient Video Frame Interpolation.** 5682-5692
- Zhiliang Wu , Changchang Sun, Hanyu Xuan, Yan Yan: **Deep Stereo Video Inpainting.** 5693-5702
- Akshay Duhane, Syed Waqas Zamir, Salman Khan, Fahad Shahbaz Khan, Ming-Hsuan Yang: **Burstormer: Burst Image Restoration and Enhancement Transformer.** 5703-5712
- Zhihang Zhong, Mingdeng Cao, Xiang Ji, Yinjiang Zheng, Imari Sato: **Blur Interpolation Transformer for Real-World Motion from Blur.** 5713-5723
- Yiheng Chi, Xingguang Zhang, Stanley H. Chan: **HDR Imaging with Spatially Varying Signal-to-Noise Ratios.** 5724-5734
- Yusaku Yoshida, Ryo Kawahara, Takahiro Okabe: **Light Source Separation and Intrinsic Image Decomposition under AC Illumination.** 5735-5743
- Yue Cao, Ming Liu, Shuai Liu, Xiaotao Wang, Lei Lei, Wangmeng Zuo: **Physics-Guided ISO-Dependent Sensor Noise Modeling for Extreme Low-Light Photography.** 5744-5753
- Yuhui Quan, Zicong Wu, Hui Ji: **Neumann Network with Recursive Kernels for Single Image Defocus Deblurring.** 5754-5763
- Carlos Rodríguez-Pardo, Henar Dominguez-Elvira, David Pascual-Hernández , Elena Garces: **UMat: Uncertainty-Aware Single Image High Resolution Material Capture.** 5764-5774
- Qingsen Yan, Song Zhang, Weiye Chen, Hao Tang, Yu Zhu, Jinqiu Sun, Luc Van Gool, Yanning Zhang: **SMAE: Few-shot Learning for HDR Deghosting with Saturation-Aware Masked Autoencoders.** 5775-5784
- Yu Zheng, Jiahui Zhan, Shengfeng He , Junyu Dong, Yong Du: **Curricular Contrastive Regularization for Physics-Aware Single Image Dehazing.** 5785-5794
- Gregory Vaksman, Michael Elad: **PatchCraft Self-Supervised Training for Correlated Image Denoising.** 5795-5804
- Miaoyu Li, Ji Liu, Ying Fu, Yulun Zhang, Dejing Dou: **Spectral Enhanced Rectangle Transformer for Hyperspectral Image Denoising.** 5805-5814
- Dongwon Park, Byung Hyun Lee, Se Young Chun: **All-in-One Image Restoration for Unknown Degradations Using Adaptive Discriminative Filters for Specific Degradations.** 5815-5824
- Jinghao Zhang, Jie Huang, Mingde Yao, Zizheng Yang, Hu Yu, Man Zhou, Feng Zhao: **Ingredient-oriented Multi-Degradation Learning for Image Restoration.** 5825-5835
- Fadi Boutros, Meiling Fang, Marcel Klemt, Biying Fu, Naser Damer: **CR-FIQA: Face Image Quality Assessment by Learning Sample Relative Classifiability.** 5836-5845
- Avinab Saha, Sandeep Mishra, Alan C. Bovik: **Re-IQA: Unsupervised Learning for Image Quality Assessment in the Wild.** 5846-5855
- Zhijun Tu, Jie Hu, Hanting Chen, Yunhe Wang: **Toward Accurate Post-Training Quantization for Image Super Resolution.** 5856-5865
- Jiacheng Li, Chang Chen, Wei Huang, Zhiqiang Lang, Fenglong Song, Youliang Yan, Zhiwei Xiong: **Learning Steerable Function for Efficient Image Resampling.** 5866-5875
- Woo Kyung Han, Byeonghun Lee, Sang Hyun Park, Kyong Hwan Jin : **ABCD : Arbitrary Bitwise Coefficient for De-Quantization.** 5876-5885
- Lingshun Kong, Jiangxin Dong, Jianjun Ge, Mingqiang Li, Jinshan Pan: **Efficient Frequency Domain-based Transformers for High-Quality Image Deblurring.** 5886-5895
- Xiang Chen, Hao Li, Mingqiang Li, Jinshan Pan: **Learning A Sparse Transformer Network for Effective Image Deraining.** 5896-5905
- Zixiang Zhao, Haowen Bai, Jiangshe Zhang, Yulun Zhang, Shuang Xu, Zudi Lin, Radu Timofte , Luc Van Gool: **CDDFuse: Correlation-Driven Dual-Branch Feature Decomposition for Multi-Modality Image Fusion.** 5906-5916
- Julian Jorge Andrade Guerreiro , Mitsuru Nakazawa, Björn Stenger: **PCT-Net: Full Resolution Image Harmonization Using Pixel-Wise Color Transformations.** 5917-5926
- Ke Wang, Michaël Gharbi, He Zhang, Zhihao Xia, Eli Shechtman: **Semi-Supervised Parametric Real-World Image Harmonization.** 5927-5936
- Chenfan Qu, Chongyu Liu, Yuliang Liu, Xinhong Chen, Dezhi Peng, Fengjun Guo, Lianwen Jin: **Towards Robust Tampered Text Detection in Document Image: New Dataset and New Solution.** 5937-5946
- Siyu Huang , Jie An, Donglai Wei, Jiebo Luo, Hanspeter Pfister: **QuantArt: Quantizing Image Style Transfer Towards High Visual Fidelity.** 5947-5956



- Takehiro Aoshima, Takashi Matsubara:
Deep Curvilinear Editing: Commutative and Nonlinear Image Manipulation for Pretrained Deep Generative Model. 5957-5967
- Ankan Kumar Bhunia, Salman H. Khan, Hisham Cholakkal, Rao Muhammad Anwer, Jorma Laaksonen , Mubarak Shah, Fahad Shahbaz Khan:
Person Image Synthesis via Denoising Diffusion Model. 5968-5976
- Gang Dai, Yifan Zhang, Qingfeng Wang, Qing Du, Zhuliang Yu, Zhuoman Liu, Shuangping Huang:
Disentangling Writer and Character Styles for Handwriting Generation. 5977-5986
- Harsh Rangwani, Lavish Bansal, Kartik Sharma, Tejan Karmali, Varun Jampani, R. Venkatesh Babu:
NoisyTwins: Class-Consistent and Diverse Image Generation Through StyleGANs. 5987-5996
- Jaskirat Singh, Stephen Gould, Liang Zheng:
High-Fidelity Guided Image Synthesis with Latent Diffusion Models. 5997-6006
- Bahjat Kawar, Shiran Zada, Oran Lang, Omer Tov, Huiwen Chang, Tali Dekel, Inbar Mosseri, Michal Irani:
Imagic: Text-Based Real Image Editing with Diffusion Models. 6007-6017
- HsiaoYuan Hsu, Xiangteng He, Yuxin Peng, Hao Kong, Qing Zhang:
PosterLayout: A New Benchmark and Approach for Content-Aware Visual-Textual Presentation Layout. 6018-6026
- Zhixing Zhang, Ligong Han, Arnab Ghosh, Dimitris N. Metaxas, Jian Ren:
SINE: SINgle Image Editing with Text-to-Image Diffusion Models. 6027-6037
- Ron Mokady, Amir Hertz, Kfir Aberman, Yael Pritch, Daniel Cohen-Or:
Null-text Inversion for Editing Real Images using Guided Diffusion Models. 6038-6047
- Gowthami Somepalli, Vasu Singla, Micah Goldblum, Jonas Geiping, Tom Goldstein:
Diffusion Art or Digital Forgery? Investigating Data Replication in Diffusion Models. 6048-6058
- Hyungjin Chung, Jeongsol Kim, Sehui Kim, Jong Chul Ye:
Parallel Diffusion Models of Operator and Image for Blind Inverse Problems. 6059-6069
- Nitin Gopalakrishnan Nair, Wele Gedara Chaminda Bandara , Vishal M. Patel:
Unite and Conquer: Plug & Play Multi-Modal Synthesis Using Diffusion Models. 6070-6079
- Ziqi Huang, Kelvin C. K. Chan, Yuming Jiang, Ziwei Liu:
Collaborative Diffusion for Multi-Modal Face Generation and Editing. 6080-6090
- Gyeongman Kim, Hajin Shim, Hyunsu Kim, Yunjey Choi, Junho Kim, Eunho Yang:
Diffusion Video Autoencoders: Toward Temporally Consistent Face Video Editing via Disentangled Video Encoding. 6091-6100
- Runsen Feng, Zongyu Guo, Weiping Li, Zhibo Chen:
NVTC: Nonlinear Vector Transform Coding. 6101-6110
- Linfeng Qi, Jiahao Li, Bin Li, Houqiang Li, Yan Lu:
Motion Information Propagation for Neural Video Compression. 6111-6120
- Xiaotao Hu, Zhewei Huang, Ailin Huang, Jun Xu, Shuchang Zhou:
A Dynamic Multi-Scale Voxel Flow Network for Video Prediction. 6121-6131
- Bo He, Xitong Yang, Hanyu Wang, Zuxuan Wu, Hao Chen, Shuaiyi Huang, Yixuan Ren , Ser-Nam Lim, Abhinav Shrivastava:
Towards Scalable Neural Representation for Diverse Videos. 6132-6142
- Shaowen Xie, Hao Zhu, Zhen Liu, Qi Zhang, You Zhou, Xun Cao, Zhan Ma:
DINER: Disorder-Invariant Implicit Neural Representation. 6143-6152
- Jiafeng Li, Ying Wen, Lianghua He:
SCConv: Spatial and Channel Reconstruction Convolution for Feature Redundancy. 6153-6162
- Xuan Shen, Yaohua Wang, Ming Lin, Yilun Huang, Hao Tang, Xiuyu Sun, Yanzhi Wang:
DeepMAD: Mathematical Architecture Design for Deep Convolutional Neural Network. 6163-6173
- Jiechong Song, Chong Mou, Shiqi Wang , Siwei Ma, Jian Zhang:
Optimization-Inspired Cross-Attention Transformer for Compressive Sensing. 6174-6184
- Ali Hassani, Steven Walton, Jiachen Li, Shen Li, Humphrey Shi:
Neighborhood Attention Transformer. 6185-6194
- Shuning Chang, Pichao Wang, Ming Lin, Fan Wang, David Junhao Zhang, Rong Jin, Mike Zheng Shou:
Making Vision Transformers Efficient from A Token Sparsification View. 6195-6205
- Gongjie Zhang, Zhipeng Luo, Zichen Tian, Jingyi Zhang, Xiaoqin Zhang, Shijian Lu:
Towards Efficient Use of Multi-Scale Features in Transformer-Based Object Detectors. 6206-6216
- Steffen Czolbe, Adrian V. Dalca:
Neuralizer: General Neuroimage Analysis without Re-Training. 6217-6230
- Saimunur Rahman, Piotr Koniusz, Lei Wang, Luping Zhou, Peyman Moghadam, Changming Sun:
Learning Partial Correlation based Deep Visual Representation for Image Classification. 6231-6240
- Xiangwen Kong, Xiangyu Zhang:
Understanding Masked Image Modeling via Learning Occlusion Invariant Feature. 6241-6251



- Jihao Liu, Xin Huang, Jinliang Zheng, Yu Liu, Hongsheng Li :
MixMAE: Mixed and Masked Autoencoder for Efficient Pretraining of Hierarchical Vision Transformers. 6252-6261
- Lai Wei, Zhengwei Chen, Jun Yin, Changming Zhu, Rigui Zhou, Jin Liu :
Adaptive Graph Convolutional Subspace Clustering. 6262-6271
- Runzhong Wang, Ziao Guo, Shaofei Jiang, Xiaokang Yang, Junchi Yan:
Deep Learning of Partial Graph Matching via Differentiable Top-K. 6272-6281
- Zhihao Lin, Yongtao Wang, Jinhe Zhang, Xiaojie Chu:
DynamicDet: A Unified Dynamic Architecture for Object Detection. 6282-6291
- Sanjoy Kundu, Sathyanarayanan N. Aakur:
IS-GGT: Iterative Scene Graph Generation with Generative Transformers. 6292-6301
- Tianlei Jin , Fangtai Guo, Qiwei Meng , Shiqiang Zhu, Xiangming Xi, Wen Wang, Zonghao Mu, Wei Song:
Fast Contextual Scene Graph Generation with Unbiased Context Augmentation. 6302-6311
- Rui Wang, Dongdong Chen, Zuxuan Wu, Yinpeng Chen, Xiyang Dai, Mengchen Liu, Lu Yuan, Yu-Gang Jiang:
Masked Video Distillation: Rethinking Masked Feature Modeling for Self-supervised Video Representation Learning. 6312-6322
- Rezaul Karim, He Zhao, Richard P. Wildes, Mennatullah Siam:
MED-VT: Multiscale Encoder-Decoder Video Transformer with Application to Object Segmentation. 6323-6333
- Richard E. L. Higgins, David F. Fouhey:
MOVES: Manipulated Objects in Video Enable Segmentation. 6334-6343
- Qihao Liu, Junfeng Wu, Yi Jiang, Xiang Bai, Alan L. Yuille, Song Bai:
InstMove: Instance Motion for Object-centric Video Segmentation. 6344-6354
- Yongqi An, Xu Zhao, Tao Yu, Haiyun Gu, Chaoyang Zhao, Ming Tang, Jinqiao Wang:
ZBS: Zero-Shot Background Subtraction via Instance-Level Background Modeling and Foreground Selection. 6355-6364
- Yiming Cui:
Feature Aggregated Queries for Transformer-Based Video Object Detectors. 6365-6376
- Anwesa Choudhuri, Girish Chowdhary, Alexander G. Schwing:
Context-Aware Relative Object Queries to Unify Video Instance and Panoptic Segmentation. 6377-6386
- Jue Wang, Wentao Zhu, Pichao Wang, Xiang Yu, Linda Liu, Mohamed Omar, Raffay Hamid:
Selective Structured State-Spaces for Long-Form Video Understanding. 6387-6397
- Xitong Yang, Fu-Jen Chu, Matt Feiszli, Raghav Goyal, Lorenzo Torresani, Du Tran:
Relational Space-Time Query in Long-Form Videos. 6398-6408
- Changan Chen, Alexander Richard, Roman Shapovalov, Vamsi Krishna Ithapu, Natalia Neverova, Kristen Grauman, Andrea Vedaldi:
Novel-View Acoustic Synthesis. 6409-6419
- Weixuan Sun, Jiayi Zhang, Jianyuan Wang, Zheyuan Liu, Yiran Zhong, Tianpeng Feng, Yandong Guo, Yanhao Zhang, Nick Barnes:
Learning Audio-Visual Source Localization via False Negative Aware Contrastive Learning. 6420-6429
- Kim Sung-Bin, Arda Senocak, Hyunwoo Ha, Andrew Owens, Tae-Hyun Oh:
Sound to Visual Scene Generation by Audio-to-Visual Latent Alignment. 6430-6440
- Junwen Xiong, Ganglai Wang, Peng Zhang, Wei Huang, Yufei Zha, Guangtao Zhai:
CASP-Net: Rethinking Video Saliency Prediction from an Audio-Visual Consistency Perceptual Perspective. 6441-6450
- Xuehao Gao, Shaoyi Du, Yang Wu, Yang Yang:
Decompose More and Aggregate Better: Two Closer Looks at Frequency Representation Learning for Human Motion Prediction. 6451-6460
- Bahar Aydemir , Ludo Hoffstetter, Tong Zhang, Mathieu Salzmann, Sabine Süsstrunk:
TempSAL - Uncovering Temporal Information for Deep Saliency Prediction. 6461-6470
- Fumiaki Sato, Ryo Hachiuma, Taiki Sekii:
Prompt-Guided Zero-Shot Anomaly Action Recognition using Pretrained Deep Skeleton Features. 6471-6480
- Xinyu Gong, Sreyas Mohan, Naina Dhingra, Jean-Charles Bazin, Yilei Li, Zhangyang Wang, Rakesh Ranjan:
MMG-Ego4D: Multi-Modal Generalization in Egocentric Action Recognition. 6481-6491
- Yuyang Wanyan, Xiaoshan Yang, Chaofan Chen, Changsheng Xu:
Active Exploration of Multimodal Complementarity for Few-Shot Action Recognition. 6492-6502
- Kaiyuan Liu, Yunheng Li, Shenglan Liu, Chenwei Tan, Zihang Shao:
Reducing the Label Bias for Timestamp Supervised Temporal Action Segmentation. 6503-6513
- Hyolim Kang, Hanjung Kim, Joungbin An, Minsu Cho, Seon Joo Kim:
Soft-Landing Strategy for Alleviating the Task Discrepancy Problem in Temporal Action Localization Tasks. 6514-6523
- Meng Cao, Fangyun Wei, Can Xu, Xiubo Geng, Long Chen , Can Zhang, Yuexian Zou, Tao Shen, Dixin Jiang:
Iterative Proposal Refinement for Weakly-Supervised Video Grounding. 6524-6534



- Shixing Chen, Chun-Hao Liu, Xiang Hao, Xiaohan Nie, Maxim Arap, Raffay Hamid:
Movies2Scenes: Using Movie Metadata to Learn Scene Representation. 6535-6544
- Hanoona Abdul Rasheed, Muhammad Uzair Khattak, Muhammad Maaz, Salman H. Khan, Fahad Shahbaz Khan:
Fine-tuned CLIP Models are Efficient Video Learners. 6545-6554
- Ruyang Liu, Jingjia Huang, Ge Li, Jiashi Feng, Xinglong Wu, Thomas H. Li:
Revisiting Temporal Modeling for CLIP-Based Image-to-Video Knowledge Transferring. 6555-6564
- Siteng Huang , Biao Gong, Yulin Pan, Jianwen Jiang, Yiliang Lv, Yuyuan Li, Donglin Wang:
VoP: Text-Video Co-Operative Prompt Tuning for Cross-Modal Retrieval. 6565-6574
- Lan Wang, Gaurav Mittal, Sandra Sajeev, Ye Yu, Matthew Hall, Vishnu Naresh Boddeti, Mei Chen:
ProTÉGé: Untrimmed Pretraining for Video Temporal Grounding by Video Temporal Grounding. 6575-6585
- Yue Zhao, Ishan Misra, Philipp Krähenbühl, Rohit Girdhar:
Learning Video Representations from Large Language Models. 6586-6597
- Jinpeng Wang, Yixiao Ge, Rui Yan, Yuying Ge, Kevin Qinghong Lin, Satoshi Tsutsui, Xudong Lin, Guanyu Cai, Jianping Wu, Ying Shan, Xiaohu Qie, Mike Zheng Shou:
All in One: Exploring Unified Video-Language Pre-Training. 6598-6608
- Chao Xu, Junwei Zhu, Jiangning Zhang, Yue Han, Wenqing Chu, Ying Tai, Chengjie Wang, Zhifeng Xie, Yong Liu:
High-Fidelity Generalized Emotional Talking Face Generation with Multi-Modal Emotion Space Learning. 6609-6619
- Wenhao Wu , Xiaohan Wang, Haipeng Luo, Jingdong Wang, Yi Yang, Wanli Ouyang:
Bidirectional Cross-Modal Knowledge Exploration for Video Recognition with Pre-trained Vision-Language Models. 6620-6630
- Yong Li, Yuanzhi Wang, Zhen Cui:
Decoupled Multimodal Distilling for Emotion Recognition. 6631-6640
- Panos Achlioptas, Maks Ovsjanikov, Leonidas J. Guibas, Sergey Tulyakov:
Affection: Learning Affective Explanations for Real-World Visual Data. 6641-6651
- Zhao Xie, Tian Gao, Kewei Wu, Jiao Chang:
An Actor-centric Causality Graph for Asynchronous Temporal Inference in Group Activity. 6652-6661
- Mengyin Liu , Jie Jiang, Chao Zhu, Xu-Cheng Yin:
VLPD: Context-Aware Pedestrian Detection via Vision-Language Semantic Self-Supervision. 6662-6671
- Jiazhao Zhang, Liu Dai, Fanpeng Meng, Qingnan Fan, Xuelin Chen, Kai Xu, He Wang:
3D-Aware Object Goal Navigation via Simultaneous Exploration and Identification. 6672-6682
- Minyoung Hwang, Jaeyeon Jeong, Minsoo Kim, Yoonseon Oh, Songhwai Oh:
Meta-Explore: Exploratory Hierarchical Vision-and-Language Navigation Using Scene Object Spectrum Grounding. 6683-6693
- Santhosh Kumar Ramakrishnan, Ziad Al-Halah, Kristen Grauman:
NaQ: Leveraging Narrations as Queries to Supervise Episodic Memory. 6694-6703
- Yao Mu, Shunyu Yao, Mingyu Ding, Ping Luo, Chuang Gan:
EC²: Emergent Communication for Embodied Control. 6704-6714
- Jingyi Xu, Tushar Vaidya, Yufei Wu, Saket Chandra, Zhangsheng Lai, Kai Fong Ernest Chong:
Abstract Visual Reasoning: An Algebraic Approach for Solving Raven's Progressive Matrices. 6715-6724
- Sergio Tascon-Morales, Pablo Márquez-Neila, Raphael Sznitman:
Logical Implications for Visual Question Answering Consistency. 6725-6735
- Shi Chen , Qi Zhao:
Divide and Conquer: Answering Questions with Object Factorization and Compositional Reasoning. 6736-6745
- Gi-Cheon Kang, Sungdong Kim, Jin-Hwa Kim, Donghyun Kwak, Byoung-Tak Zhang:
The Dialog Must Go On: Improving Visual Dialog via Generative Self-Training. 6746-6756
- Hantao Yao, Rui Zhang, Changsheng Xu:
Visual-Language Prompt Tuning with Knowledge-Guided Context Optimization. 6757-6767
- Hyeongjun Kwon, Taeyong Song, Somi Jeong, Jin Kim, Jinhyun Jang, Kwanghoon Sohn:
Probabilistic Prompt Learning for Dense Prediction. 6768-6777
- Morris Alper, Michael Fiman, Hadar Averbuch-Elor:
Is BERT Blind? Exploring the Effect of Vision-and-Language Pretraining on Visual Language Understanding. 6778-6788
- Yatai Ji, Rongcheng Tu, Jie Jiang, Weijie Kong, Chengfei Cai, Wenzhe Zhao, Hongfa Wang, Yujiu Yang , Wei Liu:
Seeing What You Miss: Vision-Language Pre-training with Semantic Completion Learning. 6789-6798
- Joya Chen, Difei Gao, Kevin Qinghong Lin, Mike Zheng Shou:
Affordance Grounding from Demonstration Video to Target Image. 6799-6808
- Hongchen Luo, Wei Zhai, Jing Zhang, Yang Cao, Dacheng Tao:
Leverage Interactive Affinity for Affordance Learning. 6809-6819
- Ashish Seth, Mayur Hemani, Chirag Agarwal:
DeAR: Debiasing Vision-Language Models with Additive Residuals. 6820-6829



- Xinlong Wang, Wen Wang, Yue Cao, Chunhua Shen, Tiejun Huang:
Images Speak in Images: A Generalist Painter for In-Context Visual Learning. 6830-6839
- Songwei Ge, Shlok Mishra, Simon Kornblith, Chun-Liang Li, David Jacobs:
Hyperbolic Contrastive Learning for Visual Representations beyond Objects. 6840-6849
- Subhadeep Koley, Ayan Kumar Bhunia, Aneeshan Sain , Pinaki Nath Chowdhury, Tao Xiang, Yi-Zhe Song :
Picture that Sketch: Photorealistic Image Generation from Abstract Sketches. 6850-6861
- Sagar Vaze, Nicolas Carion, Ishan Misra:
GeneCIS: A Benchmark for General Conditional Image Similarity. 6862-6872
- Aneeshan Sain , Ayan Kumar Bhunia, Subhadeep Koley, Pinaki Nath Chowdhury, Soumitri Chattopadhyay, Tao Xiang, Yi-Zhe Song :
Exploiting Unlabelled Photos for Stronger Fine-Grained SBIR. 6873-6883
- Chuan Tang, Xi Yang, Bojian Wu , Zhizhong Han, Yi Chang:
Parts2Words: Learning Joint Embedding of Point Clouds and Texts by Bidirectional Matching Between Parts and Words. 6884-6893
- (Withdrawn) DeltaEdit: Exploring Text-free Training for Text-Driven Image Manipulation. 6894-6903
- Rui Shao, Tianxing Wu, Ziwei Liu:
Detecting and Grounding Multi-Modal Media Manipulation. 6904-6913
- Sara Sarto , Manuele Barraco, Marcella Cornia, Lorenzo Baraldi , Rita Cucchiara:
Positive-Augmented Contrastive Learning for Image and Video Captioning Evaluation. 6914-6924
- Tal Shaharabany, Lior Wolf:
Similarity Maps for Self-Training Weakly-Supervised Phrase Grounding. 6925-6934
- Roberto Dessì, Michele Bevilacqua, Eleonora Gualdoni, Nathanaël Carraz Rakotonirina, Francesca Franzon, Marco Baroni :
Cross-Domain Image Captioning with Discriminative Finetuning. 6935-6944
- Chenhao Zheng, Ayush Shrivastava, Andrew Owens:
EXIF as Language: Learning Cross-Modal Associations between Images and Camera Metadata. 6945-6956
- Noa Garcia, Yusuke Hirota, Yankun Wu, Yuta Nakashima:
Uncurated Image-Text Datasets: Shedding Light on Demographic Bias. 6957-6966
- Filip Radenovic, Abhimanyu Dubey, Abhishek Kadian, Todor Mihaylov, Simon Vandenhende, Yash Patel, Yi Wen, Vignesh Ramanathan, Dhruv Mahajan:
Filtering, Distillation, and Hard Negatives for Vision-Language Pre-Training. 6967-6977
- Wenwen Yu, Yuliang Liu, Wei Hua, Deqiang Jiang, Bo Ren, Xiang Bai:
Turning a CLIP Model into a Scene Text Detector. 6978-6988
- Xiangjie Sui, Yuming Fang, Hanwei Zhu , Shiqi Wang, Zhou Wang:
ScanDMM: A Deep Markov Model of Scanpath Prediction for 360° Images. 6989-6999
- Thomas Stegmüller, Tim Lebailly, Behzad Bozorgtabar, Tinne Tuytelaars , Jean-Philippe Thiran:
CrOC: Cross-View Online Clustering for Dense Visual Representation Learning. 7000-7009
- Runyu Ding, Jihan Yang, Chuhui Xue, Wenqing Zhang, Song Bai, Xiaojuan Qi:
PLA: Language-Driven Open-Vocabulary 3D Scene Understanding. 7010-7019
- Runnan Chen, Youquan Liu, Lingdong Kong , Xinge Zhu, Yuexin Ma, Yikang Li, Yuenan Hou, Yu Qiao, Wenping Wang:
CLIP2Scene: Towards Label-efficient 3D Scene Understanding by CLIP. 7020-7030
- Xiaoshi Wu, Feng Zhu, Rui Zhao, Hongsheng Li :
CORA: Adapting CLIP for Open-Vocabulary Detection with Region Prompting and Anchor Pre-Matching. 7031-7040
- María Alejandra Bravo, Sudhanshu Mittal, Simon Ging, Thomas Brox:
Open-vocabulary Attribute Detection. 7041-7050
- Tao Wang :
Learning to Detect and Segment for Open Vocabulary Object Detection. 7051-7060
- Feng Liang, Bichen Wu, Xiaoliang Dai, Kunpeng Li, Yinan Zhao, Hang Zhang, Peizhao Zhang, Peter Vajda, Diana Marculescu :
Open-Vocabulary Semantic Segmentation with Mask-adapted CLIP. 7061-7070
- Muyang Yi, Quan Cui, Hao Wu, Cheng Yang, Osamu Yoshie, Hongtao Lu:
A Simple Framework for Text-Supervised Semantic Segmentation. 7071-7080
- Haoran Geng, Helin Xu, Chengyang Zhao, Chao Xu, Li Yi, Siyuan Huang, He Wang:
GAPartNet: Cross-Category Domain-Generalizable Object Perception and Manipulation via Generalizable and Actionable Parts. 7081-7091
- Chuwei Luo, Changxu Cheng, Qi Zheng, Cong Yao:
GeoLayoutLM: Geometric Pre-training for Visual Information Extraction. 7092-7101
- Anas Mahmoud, Jordan S. K. Hu, Tianshu Kuai, Ali Harakeh, Liam Paull, Steven L. Waslander:
Self-Supervised Image-to-Point Distillation via Semantically Tolerant Contrastive Loss. 7102-7110



- Jiaqi Chen, Jiachen Lu, Xiatian Zhu, Li Zhang:
Generative Semantic Segmentation. 7111-7120
- Yixuan Sun, Yiwen Huang, Haijing Guo, Yuzhou Zhao, Runmin Wu, Yizhou Yu, Weifeng Ge, Wenqiang Zhang:
MISC210K: A Large-Scale Dataset for Multi-Instance Semantic Correspondence. 7121-7130
- Yong Yang, Qiong Chen, Yuan Feng, Tianlin Huang:
MIANet: Aggregating Unbiased Instance and General Information for Few-Shot Semantic Segmentation. 7131-7140
- Vignesh Ramanathan, Anmol Kalia, Vladan Petrovic, Yi Wen, Baixue Zheng, Baishan Guo, Rui Wang, Aaron Marquez, Rama Kovvuri, Abhishek Kadian, Amir Mousavi, Yiwen Song, Abhimanyu Dubey, Dhruv Mahajan:
PACO: Parts and Attributes of Common Objects. 7141-7151
- Jang Hyun Cho, Philipp Krähenbühl, Vignesh Ramanathan:
PartDistillation: Learning Parts from Instance Segmentation. 7152-7161
- Kehan Li, Zhennan Wang, Zesen Cheng, Runyi Yu, Yian Zhao, Guoli Song, Chang Liu, Li Yuan, Jie Chen:
ACSeg: Adaptive Conceptualization for Unsupervised Semantic Segmentation. 7162-7172
- Pau de Jorge, Riccardo Volpi, Philip H. S. Torr, Grégory Rogez:
Reliability in Semantic Segmentation: Are we on the Right Track? 7173-7182
- Yuan Wang, Rui Sun , Tianzhu Zhang:
Rethinking the Correlation in Few-Shot Segmentation: A Buoys View. 7183-7192
- Ruihuang Li, Chenhang He, Yabin Zhang, Shuai Li, Liyi Chen, Lei Zhang:
SIM: Semantic-aware Instance Mask Generation for Box-Supervised Instance Segmentation. 7193-7203
- Jia-Wen Xiao, Chang-Bin Zhang, Jiekang Feng, Xiaolei Liu, Joost van de Weijer , Ming-Ming Cheng:
Endpoints Weight Fusion for Class Incremental Semantic Segmentation. 7204-7213
- Chao Shang, Hongliang Li, Fanman Meng, Qingbo Wu, Heqian Qiu, Lanxiao Wang:
Incrementer: Transformer for Class-Incremental Semantic Segmentation with Knowledge Distillation Focusing on Old Class. 7214-7224
- Rui Gong, Qin Wang, Martin Danelljan, Dengxin Dai, Luc Van Gool:
Continuous Pseudo-Label Rectified Domain Adaptive Semantic Segmentation with Implicit Neural Representations. 7225-7235
- Lihe Yang, Lei Qi, Litong Feng, Wayne Zhang , Yinghuan Shi:
Revisiting Weak-to-Strong Consistency in Semi-Supervised Semantic Segmentation. 7236-7246
- Long Li , Junwei Han, Ni Zhang, Nian Liu, Salman H. Khan, Hisham Cholakkal, Rao Muhammad Anwer, Fahad Shahbaz Khan:
Discriminative Co-Saliency and Background Mining Transformer for Co-Salient Object Detection. 7247-7256
- Huajun Zhou , Bo Qiao, Lingxiao Yang, Jianhuang Lai, Xiaohua Xie:
Texture-Guided Saliency Distilling for Unsupervised Salient Object Detection. 7257-7267
- Dongliang Chang, Yujun Tong, Ruoyi Du, Timothy M. Hospedales, Yi-Zhe Song, Zhanyu Ma:
An Erudite Fine-Grained Visual Classification Model. 7268-7277
- Yuan Wang, Kun Yu, Chen Chen, Xiyuan Hu, Silong Peng:
Dynamic Graph Learning with Content-guided Spatial-Frequency Relation Reasoning for Deepfake Detection. 7278-7287
- Yanbei Chen, Manchen Wang, Abhay Mittal, Zhenlin Xu, Paolo Favaro, Joseph Tighe, Davide Modolo:
ScaleDet: A Scalable Multi-Dataset Object Detector. 7288-7297
- Tenghao Cai, Zhizhong Zhang, Xin Tan, Yanyun Qu, Guannan Jiang, Chengjie Wang, Yuan Xie:
Multi-Centroid Task Descriptor for Dynamic Class Incremental Inference. 7298-7307
- Min Shi, Zihao Huang, Xianzheng Ma, Xiaowei Hu, Zhiguo Cao:
Matching Is Not Enough: A Two-Stage Framework for Category-Agnostic Pose Estimation. 7308-7317
- Chang Xu, Jian Ding, Jinwang Wang, Wen Yang, Huai Yu, Lei Yu, Gui-Song Xia :
Dynamic Coarse-to-Fine Learning for Oriented Tiny Object Detection. 7318-7328
- Shilong Zhang, Xinjiang Wang, Jiaqi Wang, Jiangmiao Pang, Chengqi Lyu, Wenwei Zhang, Ping Luo, Kai Chen:
Dense Distinct Query for End-to-End Object Detection. 7329-7338
- Berkan Demirel, Orhun Bugra Baran, Ramazan Gokberk Cinbis:
Meta-Tuning Loss Functions and Data Augmentation for Few-Shot Object Detection. 7339-7349
- Shuai Li, Minghan Li , Ruihuang Li, Chenhang He, Lei Zhang:
One-to-Few Label Assignment for End-to-End Dense Detection. 7350-7359
- Olga Veksler:
Test Time Adaptation with Regularized Loss for Weakly Supervised Salient Object Detection. 7360-7369
- Liang Liu, Boshen Zhang, Jiangning Zhang, Wuhao Zhang, Zhenye Gan, Guanzhong Tian, Wenbing Zhu, Yabiao Wang, Chengjie Wang:
MixTeacher: Mining Promising Labels with Mixed Scale Teacher for Semi-Supervised Object Detection. 7370-7379
- Yunqing Zhao, Chao Du, Milad Abdollahzadeh, Tianyu Pang , Min Lin, Shuicheng Yan, Ngai-Man Cheung:
Exploring Incompatible Knowledge Transfer in Few-shot Image Generation. 7380-7391



Yunfei Zhang, Xiaoyang Huo, Tianyi Chen, Si Wu, Hau-San Wong:
Exploring Intra-class Variation Factors with Learnable Cluster Prompts for Semi-supervised Image Synthesis.

7392-7401

Xiaoyu Liu, Bo Hu, Mingxing Li, Wei Huang, Yueyi Zhang, Zhiwei Xiong:
A Soma Segmentation Benchmark in Full Adult Fly Brain. 7402-7411

Hyungseob Shin, Hyeongyu Kim, Sewon Kim, Yohan Jun, Taejoon Eo, Dosik Hwang:
SDC-UDA: Volumetric Unsupervised Domain Adaptation Framework for Slice-Direction Continuous Cross-Modality Medical Image Segmentation. 7412-7421

Qixin Hu, Yixiong Chen, Junfei Xiao, Shuwen Sun, Jieneng Chen, Alan L. Yuille, Zongwei Zhou:
Label-Free Liver Tumor Segmentation. 7422-7432

Tim Tanida, Philip Müller, Georgios Kaassis, Daniel Rueckert:
Interactive and Explainable Region-guided Radiology Report Generation. 7433-7442

Shengxuming Zhang, Tianqi Shi, Yang Jiang, Xiuming Zhang, Jie Lei, Zunlei Feng, Mingli Song:
A Loopback Network for Explainable Microvascular Invasion Classification. 7443-7453

Honglin Li, Chenglu Zhu, Yunlong Zhang, Yuxuan Sun, Zhongyi Shui, Wenwei Kuang, Sunyi Zheng, Lin Yang:
Task-Specific Fine-Tuning via Variational Information Bottleneck for Weakly-Supervised Pathology Whole Slide Image Classification. 7454-7463

Chien-Yao Wang, Alexey Bochkovskiy, Hong-Yuan Mark Liao:
YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors. 7464-7475

Takumi Kobayashi:
Two-Way Multi-Label Loss. 7476-7485

Matthew Walmer, Saksham Suri, Kamal Gupta, Abhinav Shrivastava:
Teaching Matters: Investigating the Role of Supervision in Vision Transformers. 7486-7496

Qinghai Zheng, Jihua Zhu, Haoyu Tang:
Label Information Bottleneck for Label Enhancement. 7497-7506

Haoyu Wang , Guansong Pang , Peng Wang, Lei Zhang, Wei Wei, Yanning Zhang:
Glocal Energy-based Learning for Few-Shot Open-Set Recognition. 7507-7516

Haochen Han, Kaiyao Miao, Qinghua Zheng, Minnan Luo:
Noisy Correspondence Learning with Meta Similarity Correction. 7517-7526

Daniel J. Trosten , Riddhi Chakraborty, Sigurd Løkse, Kristoffer Knutsen Wickstrøm, Robert Jenssen, Michael C. Kampffmeyer:
Hubs and Hyperspheres: Reducing Hubness and Improving Transductive Few-Shot Learning with Hyperspherical Embeddings. 7527-7536

Sungnyun Kim, Sangmin Bae, Se-Young Yun:
Coreset Sampling from Open-Set for Fine-Grained Self-Supervised Learning. 7537-7547

Yuhao Chen, Xin Tan, Borui Zhao, Zhaowei Chen, Renjie Song, Jiajun Liang, Xuequan Lu :
Boosting Semi-Supervised Learning by Exploiting All Unlabeled Data. 7548-7557

Yanxi Li, Chang Xu:
Trade-off between Robustness and Accuracy of Vision Transformers. 7558-7568

Shibin Mei, Chenglong Zhao, Shengchao Yuan, Bingbing Ni:
Exploring and Utilizing Pattern Imbalance. 7569-7578

Nan Pu, Zhun Zhong, Nicu Sebe :
Dynamic Conceptual Contrastive Learning for Generalized Category Discovery. 7579-7588

Miguel Á. Carreira-Perpiñán, Magzhan Gabdolla, Arman Zharmagambetov:
Towards Better Decision Forests: Forest Alternating Optimization. 7589-7598

Yi-Kai Zhang, Qi-Wei Wang, De-Chuan Zhan, Han-Jia Ye:
Learning Debiased Representations via Conditional Attribute Interpolation. 7599-7608

Deng-Bao Wang, Lanqing Li, Peilin Zhao, Pheng-Ann Heng, Min-Ling Zhang:
On the Pitfall of Mixup for Uncertainty Calibration. 7609-7618

Yixin Zhang, Zilei Wang, Weinan He:
Class Relationship Embedded Learning for Source-Free Unsupervised Domain Adaptation. 7619-7629

Xinjiang Wang, Zeyu Liu, Yu Hu, Wei Xi, Wenxian Yu, Danping Zou:
FeatureBooster: Boosting Feature Descriptors with a Lightweight Neural Network. 7630-7639

Mattia Litrico , Alessio Del Bue , Pietro Mororio:
Guiding Pseudo-labels with Uncertainty Estimation for Source-free Unsupervised Domain Adaptation. 7640-7650

Duojun Huang, Jichang Li, Weikai Chen, Junshi Huang, Zhenhua Chai, Guanbin Li:
Divide and Adapt: Active Domain Adaptation via Customized Learning. 7651-7660

Qian Jiang, Changyou Chen, Han Zhao, Liqun Chen, Qing Ping, Son Dinh Tran, Yi Xu, Belinda Zeng, Trishul Chilimbi:
Understanding and Constructing Latent Modality Structures in Multi-Modal Representation Learning. 7661-7671

Chengkun Wang, Wenzhao Zheng, Junlong Li, Jie Zhou, Jiwen Lu:
Deep Factorized Metric Learning. 7672-7682



- Jin Chen, Zhi Gao, Xinxiao Wu, Jiebo Luo:
Meta-Causal Learning for Single Domain Generalization. 7683-7692
- Ondrej Bohdal, Yinbing Tian, Yongshuo Zong, Ruchika Chavhan, Da Li, Henry Gouk, Li Guo, Timothy M. Hospedales:
Meta Omnia: A Benchmark for General-Purpose Learning-to-Learn. 7693-7703
- Mario Döbler, Robert A. Marsden, Bin Yang:
Robust Mean Teacher for Continual and Gradual Test-Time Adaptation. 7704-7714
- Yun Yi, Haokui Zhang, Wenze Hu, Nannan Wang, Xiaoyu Wang:
NAR-Former: Neural Architecture Representation Learning Towards Holistic Attributes Prediction. 7715-7724
- Cheng-Hao Tu, Zheda Mai, Wei-Lun Chao:
Visual Query Tuning: Towards Effective Usage of Intermediate Representations for Parameter and Memory Efficient Transfer Learning. 7725-7735
- Zixuan Hu, Li Shen , Zhenyi Wang, Tongliang Liu, Chun Yuan, Dacheng Tao:
Architecture, Dataset and Model-Scale Agnostic Data-free Meta-Learning. 7736-7745
- Huijing Zhuang, Zhenyu Weng, Run He, Zhiping Lin, Ziqian Zeng:
GKEAL: Gaussian Kernel Embedded Analytic Learning for Few-Shot Class Incremental Task. 7746-7755
- Chuntao Ding, Zhichao Lu , Shangguang Wang, Ran Cheng, Vishnu Naresh Boddeti:
Mitigating Task Interference in Multi-Task Learning via Explicit Task Routing with Non-Learnable Primitives. 7756-7765
- Min Chen, Weizhuo Gao, Gaoyang Liu, Kai Peng, Chen Wang :
Boundary Unlearning: Rapid Forgetting of Deep Networks via Shifting the Decision Boundary. 7766-7775
- Wenjin Wang, Yunqing Hu, Qianglong Chen, Yin Zhang:
Task Difficulty Aware Parameter Allocation & Regularization for Lifelong Learning. 7776-7785
- Gaurav Patel, Konda Reddy Mopuri, Qiang Qiu:
Learning to Retain while Acquiring: Combating Distribution-Shift in Adversarial Data-Free Knowledge Distillation. 7786-7794
- Yizhuo Chen, Kaizhao Liang, Zhe Zeng, Shuochao Yao, Huajie Shao:
A Unified Knowledge Distillation Framework for Deep Directed Graphical Models. 7795-7804
- Jimuyang Zhang, Zanming Huang, Eshed Ohn-Bar:
Coaching a Teachable Student. 7805-7815
- Yan-Shuo Liang, Wu-Jun Li:
Adaptive Plasticity Improvement for Continual Learning. 7816-7825
- Lianzhe Wang, Shiji Zhou, Shanghang Zhang, Xu Chu, Heng Chang, Wenwu Zhu:
Improving Generalization of Meta-Learning with Inverted Regularization at Inner-Level. 7826-7835
- Junjiao Tian, Xiaoliang Dai, Chih-Yao Ma, Zecheng He, Yen-Cheng Liu, Zsolt Kira:
Trainable Projected Gradient Method for Robust Fine-Tuning. 7836-7845
- Siwei Chen, Xiao Ma, Zhongwen Xu:
Imitation Learning as State Matching via Differentiable Physics. 7846-7855
- Ganlong Zhao, Guanbin Li, Yipeng Qin, Yizhou Yu:
Improved Distribution Matching for Dataset Condensation. 7856-7865
- Hongwei Yong, Ying Sun, Lei Zhang:
A General Regret Bound of Preconditioned Gradient Method for DNN Training. 7866-7875
- Jie Chen, Zilong Li, Yin Zhu, Junping Zhang, Jian Pu:
From Node Interaction to Hop Interaction: New Effective and Scalable Graph Learning Paradigm. 7876-7885
- Qi Xu, Yixin Li, Jiangrong Shen, Jian K. Liu , Huajin Tang, Gang Pan:
Constructing Deep Spiking Neural Networks from Artificial Neural Networks with Knowledge Distillation. 7886-7895
- Tong Bu, Jianhao Ding, Zecheng Hao, Zhaofei Yu:
Rate Gradient Approximation Attack Threats Deep Spiking Neural Networks. 7896-7906
- Pavan Kumar Anasosalu Vasu, James Gabriel, Jeff Zhu, Oncel Tuzel, Anurag Ranjan:
MobileOne: An Improved One millisecond Mobile Backbone. 7907-7917
- Lingjing Kong, Martin Q. Ma, Guangyi Chen , Eric P. Xing, Yuejie Chi, Louis-Philippe Morency, Kun Zhang:
Understanding Masked Autoencoders via Hierarchical Latent Variable Models. 7918-7928
- Geon Yeong Park, Sangmin Lee, Sang Wan Lee, Jong Chul Ye:
Training Debiased Subnetworks with Contrastive Weight Pruning. 7929-7938
- Ivan Koryakovskiy, Alexandra Yakovleva, Valentin Buchnev, Temur Isaev, Gleb Odinokikh:
One-Shot Model for Mixed-Precision Quantization. 7939-7949
- Yuexiao Ma, Huixia Li, Xiawu Zheng, Xuefeng Xiao, Rui Wang, Shilei Wen, Xin Pan, Fei Chao, Rongrong Ji:
Solving Oscillation Problem in Post-Training Quantization Through a Theoretical Perspective. 7950-7959
- Biao Qian, Yang Wang, Richang Hong, Meng Wang:
Adaptive Data-Free Quantization. 7960-7968



- Zheng Xu, Maxwell D. Collins, Yuxiao Wang, Liviu Panait, Sewoong Oh, Sean Augenstein, Ting Liu, Florian Schroff, H. Brendan McMahan:
Learning to Generate Image Embeddings with User-Level Differential Privacy. 7969-7980
- Matthew L. Olson, Shusen Liu, Rushil Anirudh, Jayaraman J. Thiagarajan, Peer-Timo Bremer, Weng-Keen Wong:
Cross-GAN Auditing: Unsupervised Identification of Attribute Level Similarities and Differences Between Pretrained Generative Models. 7981-7990
- Austin Xu, Mariya I. Vasileva, Achal Dave, Arjun Seshadri:
HandsOff: Labeled Dataset Generation With No Additional Human Annotations. 7991-8000
- Simone Barattin, Christos Tzelepis , Ioannis Patras, Nicu Sebe :
Attribute-Preserving Face Dataset Anonymization via Latent Code Optimization. 8001-8010
- Mert Bülent Sarıyıldız, Karteek Alahari, Diane Larlus, Yannis Kalantidis:
Fake it Till You Make it: Learning Transferable Representations from Synthetic ImageNet Clones. 8011-8021
- Hui Lv, Zhongqi Yue, Qianru Sun, Bin Luo, Zhen Cui, Hanwang Zhang:
Unbiased Multiple Instance Learning for Weakly Supervised Video Anomaly Detection. 8022-8031
- Yue Wang, Jinlong Peng, Jiangning Zhang, Ran Yi, Yabiao Wang, Chengjie Wang:
Multimodal Industrial Anomaly Detection via Hybrid Fusion. 8032-8041
- Jiaxu Miao, Zongxin Yang, Leilei Fan, Yi Yang:
FedSeg: Class-Heterogeneous Federated Learning for Semantic Segmentation. 8042-8052
- Andrey Zhmoginov, Mark Sandler, Nolan Miller, Gus Kristiansen, Max Vladymyrov:
Decentralized Learning with Multi-Headed Distillation. 8053-8063
- Chun-Mei Feng, Bangjun Li , Xinxing Xu, Yong Liu, Huazhu Fu , Wangmeng Zuo:
Learning Federated Visual Prompt in Null Space for MRI Reconstruction. 8064-8073
- Jian-Hui Duan, Wenzhong Li, Derun Zou, Ruichen Li, Sanglu Lu:
Federated Learning with Data-Agnostic Distribution Fusion. 8074-8083
- Nurbek Tastan, Karthik Nandakumar:
CaPriDe Learning: Confidential and Private Decentralized Learning Based on Encryption-Friendly Distillation Loss. 8084-8092
- Mingjun Xu, Lingyun Qin, Weijie Chen, Shiliang Pu, Lei Zhang:
Multi-view Adversarial Discriminator: Mine the Non-causal Factors for Object Detection in Unseen Domains. 8103-8112
- Mingjie Sun, Zico Kolter:
Single Image Backdoor Inversion via Robust Smoothed Classifiers. 8113-8122
- Yiming Chen, Jinyu Tian, Xiangyu Chen, Jiantao Zhou:
Effective Ambiguity Attack Against Passport-based DNN Intellectual Property Protection Schemes through Fully Connected Layer Substitution. 8123-8132
- Wenbo Jiang, Hongwei Li, Guowen Xu, Tianwei Zhang:
Color Backdoor: A Robust Poisoning Attack in Color Space. 8133-8142
- Beini Xie, Heng Chang, Ziwei Zhang, Xin Wang, Daixin Wang, Zhiqiang Zhang, Rex Ying, Wenwu Zhu:
Adversarially Robust Neural Architecture Search for Graph Neural Networks. 8143-8152
- Anqi Zhao, Tong Chu, Yahao Liu, Wen Li, Jingjing Li, Lixin Duan:
Minimizing Maximum Model Discrepancy for Transferable Black-box Targeted Attacks. 8153-8162
- Kaisheng Liang, Bin Xiao:
StyLess: Boosting the Transferability of Adversarial Examples. 8163-8172
- Jianping Zhang, Jen-tse Huang , Wenxuan Wang, Yichen Li, Weibin Wu, Xiaosen Wang, Yuxin Su, Michael R. Lyu:
Improving the Transferability of Adversarial Samples by Path-Augmented Method. 8173-8182
- Woo Jae Kim, Yoonki Cho, Junsik Jung, Sung-Eui Yoon:
Feature Separation and Recalibration for Adversarial Robustness. 8183-8192
- Zeming Wei, Yifei Wang, Yiwen Guo, Yisen Wang:
CFA: Class-Wise Calibrated Fair Adversarial Training. 8193-8201
- Shihua Huang, Zhichao Lu, Kalyanmoy Deb, Vishnu Naresh Boddeti:
Revisiting Residual Networks for Adversarial Robustness. 8202-8211
- Zhibo Wang, He Wang, Shuaifan Jin, Wenwen Zhang, Jiahui Hu, Yan Wang, Peng Sun, Wei Yuan, Kaixin Liu, Kui Ren:
Privacy-preserving Adversarial Facial Features. 8212-8221
- Dong Li , Jiaying Zhu, Menglu Wang, Jiawei Liu, Xueyang Fu, Zheng-Jun Zha:
Edge-aware Regional Message Passing Controller for Image Forgery Localization. 8222-8232
- Alankar Kotwal, Anat Levin, Ioannis Gkioulekas :
Swept-Angle Synthetic Wavelength Interferometry. 8233-8243
- Xudong Huang, Wei Li, Jie Hu, Hanting Chen, Yunhe Wang:
RefSR-NeRF: Towards High Fidelity and Super Resolution View Synthesis. 8244-8253
- Jiawei Yang, Marco Pavone , Yue Wang:
FreeNeRF: Improving Few-Shot Neural Rendering with Free Frequency Regularization. 8254-8263



- Yue Chen, Xingyu Chen, Xuan Wang, Qi Zhang, Yu Guo, Ying Shan, Fei Wang:
Local-to-Global Registration for Bundle-Adjusting Neural Radiance Fields. 8264-8273
- Xiaoshuai Zhang, Abhijit Kundu , Thomas A. Funkhouser, Leonidas J. Guibas, Hao Su, Kyle Genova:
Nerflets: Local Radiance Fields for Efficient Structure-Aware 3D Scene Representation from 2D Supervision. 8274-8284
- Zhiwen Yan, Chen Li, Gim Hee Lee:
NeRF-DS: Neural Radiance Fields for Dynamic Specular Objects. 8285-8295
- Lining Xu, Yuanbo Xiangli, Sida Peng, Xingang Pan, Nanxuan Zhao, Christian Theobalt, Bo Dai, Dahua Lin:
Grid-guided Neural Radiance Fields for Large Urban Scenes. 8296-8306
- Ziyu Wan, Christian Richardt, Aljaz Bozic, Chao Li, Vijay Rengarajan, Seonghyeon Nam, Xiaoyu Xiang, Tuotuo Li, Bo Zhu, Rakesh Ranjan, Jing Liao :
Learning Neural Duplex Radiance Fields for Real-Time View Synthesis. 8307-8316
- Chengwei Zheng, Wenbin Lin, Feng Xu:
EditableNeRF: Editing Topologically Varying Neural Radiance Fields by Key Points. 8317-8327
- Junli Cao, Huan Wang, Pavlo Chemerys, Vladislav Shakhrai, Ju Hu, Yun Fu, Denys Makoviichuk, Sergey Tulyakov, Jian Ren:
Real-Time Neural Light Field on Mobile Devices. 8328-8337
- Kunhao Liu, Fangneng Zhan, Yiwen Chen, Jiahui Zhang, Yingchen Yu, Abdulmotaleb El-Saddik, Shijian Lu, Eric P. Xing:
StyleRF: Zero-Shot 3D Style Transfer of Neural Radiance Fields. 8338-8348
- Tao Hu, Xiaogang Xu, Shu Liu, Jiaya Jia:
Point2Pix: Photo-Realistic Point Cloud Rendering via Neural Radiance Fields. 8349-8358
- Jen-Hao Rick Chang, Wei-Yu Chen, Anurag Ranjan, Kwang Moo Yi, Oncel Tuzel:
Pointersect: Neural Rendering with Cloud-Ray Intersection. 8359-8369
- Zian Wang, Tianchang Shen, Jun Gao, Shengyu Huang, Jacob Munkberg, Jon Hasselgren, Zan Gojcic, Wenzheng Chen, Sanja Fidler:
Neural Fields Meet Explicit Geometric Representations for Inverse Rendering of Urban Scenes. 8370-8380
- Zongrui Li , Qian Zheng, Boxin Shi, Gang Pan, Xudong Jiang :
DANI-Net: Uncalibrated Photometric Stereo by Differentiable Shadow Handling, Anisotropic Reflectance Modeling, and Neural Inverse Rendering. 8381-8391
- Junyoung Choi, SeokYeong Lee , Haesol Park, Seung-Won Jung, Ig-Jae Kim, Junghyun Cho:
MAIR: Multi-View Attention Inverse Rendering with 3D Spatially-Varying Lighting Estimation. 8392-8401
- Renjiao Yi, Chenyang Zhu, Kai Xu:
Weakly-supervised Single-view Image Relighting. 8402-8411
- David Futschik, Kelvin Ritland, James Vecore, Sean Fanello, Sergio Orts-Escalano, Brian Curless, Daniel Sýkora, Rohit Pandey:
Controllable Light Diffusion for Portraits. 8412-8421
- Jiabao Lei, Jiapeng Tang, Kui Jia:
RGBD2: Generative Scene Synthesis via Incremental View Inpainting Using RGBD Diffusion Models. 8422-8434
- Wenqi Xian, Aljaz Bozic, Noah Snavely, Christoph Lassner:
Neural Lens Modeling. 8435-8445
- Luke Melas-Kyriazi, Iro Laina, Christian Rupprecht, Andrea Vedaldi:
RealFusion 360° Reconstruction of Any Object from a Single Image. 8446-8455
- Zhaoshuo Li, Thomas Müller, Alex Evans, Russell H. Taylor, Mathias Unberath, Ming-Yu Liu, Chen-Hsuan Lin:
Neuralangelo: High-Fidelity Neural Surface Reconstruction. 8456-8465
- Radu Alexandru Rosu, Sven Behnke:
PermutoSDF: Fast Multi-View Reconstruction with Implicit Surfaces Using Permutohedral Lattices. 8466-8475
- Bowen Cai, Jinchi Huang, Rongfei Jia, Chengfei Lv, Huan Fu :
NeuDA: Neural Deformable Anchor for High-Fidelity Implicit Surface Reconstruction. 8476-8485
- Yunfan Ye, Renjiao Yi, Zhirui Gao, Chenyang Zhu, Zhiping Cai, Kai Xu:
NEF: Neural Edge Fields for 3D Parametric Curve Reconstruction from Multi-View Images. 8486-8495
- Seung Wook Kim, Bradley Brown, Kangxue Yin, Karsten Kreis, Katja Schwarz, Daqing Li, Robin Rombach, Antonio Torralba, Sanja Fidler:
NeuralField-LDM: Scene Generation with Hierarchical Latent Diffusion Models. 8496-8506
- Minjung Son , Jeong Joon Park , Leonidas J. Guibas, Gordon Wetzstein:
SinGRAF: Learning a 3D Generative Radiance Field for a Single Scene. 8507-8517
- Shangzhan Zhang, Sida Peng, Tianrun Chen, Linzhan Mou, Haotong Lin, Kaicheng Yu, Yiyi Liao, Xiaowei Zhou:
Painting 3D Nature in 2D: View Synthesis of Natural Scenes from a Single Semantic Mask. 8518-8528
- Hoseok Do, Eunkyoung Yoo, Taehyeong Kim, Chul Lee, Jin Young Choi:
Quantitative Manipulation of Custom Attributes on 3D-Aware Image Synthesis. 8529-8538
- Yu Yin, Kamran Ghasedi, HsiangTao Wu, Jiaolong Yang , Xin Tong, Yun Fu:
NeRFInvertor: High Fidelity NeRF-GAN Inversion for Single-Shot Real Image Animation. 8539-8548



- Jianhui Li, Jianmin Li, Haoji Zhang, Shilong Liu, Zhengyi Wang, Zihao Xiao, Kaiwen Zheng, Jun Zhu: **PREIM3D: 3D Consistent Precise Image Attribute Editing from a Single Image.** 8549-8558
- Xianghao Xu, Paul Guerrero, Matthew Fisher, Siddhartha Chaudhuri, Daniel Ritchie: **Unsupervised 3D Shape Reconstruction by Part Retrieval and Assembly.** 8559-8567
- Wenliang Zhao, Yongming Rao, Weikang Shi, Zuyan Liu, Jie Zhou, Jiwen Lu: **DiffSwap: High-Fidelity and Controllable Face Swapping via 3D-Aware Masked Diffusion.** 8568-8577
- Zhian Liu, Maomao Li, Yong Zhang, Cairong Wang, Qi Zhang, Jue Wang, Yongwei Nie: **Fine-Grained Face Swapping Via Regional GAN Inversion.** 8578-8587
- Haiyu Wu, Grace Bezold, Aman Bhatta, Kevin W. Bowyer: **Logical Consistency and Greater Descriptive Power for Facial Hair Attribute Learning.** 8588-8597
- Yuxuan Han, Zhibo Wang, Feng Xu: **Learning a 3D Morphable Face Reflectance Model from Low-Cost Data.** 8598-8608
- Sasikarn Khwanmuang, Pakkapon Phongthawee, Patsorn Sangkloy, Supasorn Suwajanakorn: **StyleGAN Salon: Multi-View Latent Optimization for Pose-Invariant Hairstyle Transfer.** 8609-8618
- Anurag Ranjan, Kwang Moo Yi, Jen-Hao Rick Chang, Oncel Tuzel: **FaceLit: Neural 3D Relightable Faces.** 8619-8628
- Alexandros Lattas, Stylianos Moschoglou, Stylianos Ploumpis, Baris Gecer, Jiankang Deng, Stefanos Zafeiriou: **FitMe: Deep Photorealistic 3D Morphable Model Avatars.** 8629-8640
- Ziyan Wang, Giljoo Nam, Tuur Stuyck, Stephen Lombardi, Chen Cao, Jason M. Saragih, Michael Zollhöfer, Jessica K. Hodgins, Christoph Lassner: **NeuWigs: A Neural Dynamic Model for Volumetric Hair Capture and Animation.** 8641-8651
- Wenxuan Zhang, Xiaodong Cun, Xuan Wang, Yong Zhang, Xi Shen, Yu Guo, Ying Shan, Fei Wang: **SadTalker: Learning Realistic 3D Motion Coefficients for Stylized Audio-Driven Single Image Talking Face Animation.** 8652-8661
- Tingting Liao, Xiaomei Zhang, Yuliang Xiu, Hongwei Yi, Xudong Liu, Guo-Jun Qi, Yong Zhang, Xuan Wang, Xiangyu Zhu, Zhen Lei: **High-Fidelity Clothed Avatar Reconstruction from a Single Image.** 8662-8672
- Nhat Le, Trong-Thang Pham, Tuong Do, Erman Tjiputra, Quang D. Tran , Anh Nguyen: **Music-Driven Group Choreography.** 8673-8682
- Xingyu Chen, Baoyuan Wang, Heung-Yeung Shum: **Hand Avatar: Free-Pose Hand Animation and Rendering from Monocular Video.** 8683-8693
- Zijun Cui, Chenyi Kuang, Tian Gao, Kartik Talamadupula, Qiang Ji: **Biomechanics-Guided Facial Action Unit Detection Through Force Modeling.** 8694-8703
- Jiashun Wang, Xuetong Li, Sifei Liu, Shalini De Mello, Orazio Gallo, Xiaolong Wang, Jan Kautz: **Zero-shot Pose Transfer for Unrigged Stylized 3D Characters.** 8704-8714
- Yash Kant, Aliaksandr Siarohin, Riza Alp Güler, Menglei Chai, Jian Ren, Sergey Tulyakov, Igor Gilitschenski: **Invertible Neural Skinning.** 8715-8725
- Michael J. Black, Priyanka Patel, Joachim Tesch, Jinlong Yang: **BEDLAM: A Synthetic Dataset of Bodies Exhibiting Detailed Lifelike Animated Motion.** 8726-8737
- Dae-Young Song, HeeKyung Lee, Jeongil Seo, Donghyeon Cho: **DIFu: Depth-Guided Implicit Function for Clothed Human Reconstruction.** 8738-8747
- Junying Wang, Jae Shin Yoon, Tuanfeng Y. Wang, Krishna Kumar Singh, Ulrich Neumann: **Complete 3D Human Reconstruction from a Single Incomplete Image.** 8748-8758
- Chen Geng , Sida Peng, Zhen Xu, Hujun Bao, Xiaowei Zhou: **Learning Neural Volumetric Representations of Dynamic Humans in Minutes.** 8759-8770
- Weixiao Liu, Yuwei Wu, Sipu Ruan, Gregory S. Chirikjian: **Marching-Primitives: Shape Abstraction from Signed Distance Function.** 8771-8780
- Qi Fang, Kang Chen, Yinghui Fan, Qing Shuai, Jiefeng Li, Weidong Zhang: **Learning Analytical Posterior Probability for Human Mesh Recovery.** 8781-8791
- Shangzhe Wu, Ruining Li, Tomas Jakab, Christian Rupprecht, Andrea Vedaldi: **MagicPony: Learning Articulated 3D Animals in the Wild.** 8792-8802
- Wenqiang Xu, Zhenjun Yu, Han Xue, Ruolin Ye, Siqiong Yao, Cewu Lu: **Visual-Tactile Sensing for In-Hand Object Reconstruction.** 8803-8812
- Ruihang Chu, Zhengze Liu, Xiaoqing Ye, Xiao Tan, Xiaojuan Qi, Chi-Wing Fu, Jiaya Jia: **Command-driven Articulated Object Understanding and Manipulation.** 8813-8823
- Jirong Liu, Ruo Zhang, Haoshu Fang, Minghao Gou, Hongjie Fang, Chenxi Wang, Sheng Xu, Hengxu Yan, Cewu Lu: **Target-referenced Reactive Grasping for Dynamic Objects.** 8824-8833
- Juze Zhang, Haimin Luo, Hongdi Yang, Xinru Xu, Qianyang Wu, Ye Shi, Jingyi Yu, Lan Xu, Jingya Wang: **NeuralDome: A Neural Modeling Pipeline on Multi-View Human-Object Interactions.** 8834-8845



- Changlong Jiang, Yang Xiao, Cunlin Wu, Mingyang Zhang, Jinghong Zheng, Zhiguo Cao, Joey Tianyi Zhou: **A2J-Transformer: Anchor-to-Joint Transformer Network for 3D Interacting Hand Pose Estimation from a Single RGB Image.** 8846-8855
- Yu Sun , Qian Bao, Wu Liu, Tao Mei, Michael J. Black: **TRACE: 5D Temporal Regression of Avatars with Dynamic Cameras in 3D Environments.** 8856-8866
- Nadine Rüegg, Shashank Tripathi, Konrad Schindler, Michael J. Black, Silvia Zuffi: **BITE: Beyond Priors for Improved Three-D Dog Pose Estimation.** 8867-8876
- Qitao Zhao, Ce Zheng, Mengyuan Liu, Pichao Wang, Chen Chen: **PoseFormerV2: Exploring Frequency Domain for Efficient and Robust 3D Human Pose Estimation.** 8877-8886
- Xiaolong Shen, Zongxin Yang, Xiaohan Wang, Jianxin Ma, Chang Zhou, Yi Yang: **Global-to-Local Modeling for Video-Based 3D Human Pose and Shape Estimation.** 8887-8896
- Cheng Zhang , Hai Liu, Yongjian Deng, Bochen Xie , Youfu Li : **TokenHPE: Learning Orientation Tokens for Efficient Head Pose Estimation via Transformers.** 8897-8906
- Zhengxi Hu, Yuxue Yang, Xiaolin Zhai, Dingye Yang, Bohan Zhou, Jingtai Liu: **GFIE: A Dataset and Baseline for Gaze-Following from 2D to 3D in Indoor Environments.** 8907-8916
- Yang Tian, Jiyao Zhang, Zekai Yin, Hao Dong: **Robot Structure Prior Guided Temporal Attention for Camera-to-Robot Pose Estimation from Image Sequence.** 8917-8926
- Yang Hai, Rui Song, Jiaojiao Li, Mathieu Salzmann, Yinlin Hu: **Rigidity-Aware Detection for 6D Object Pose Estimation.** 8927-8936
- Hao Wen, Jing Huang, Huili Cui, Haozhe Lin, Yu-Kun Lai, Lu Fang, Kun Li: **Crowd3D: Towards Hundreds of People Reconstruction from a Single Image.** 8937-8946
- Heng Yang, Marco Pavone: **Object Pose Estimation with Statistical Guarantees: Conformal Keypoint Detection and Geometric Uncertainty Propagation.** 8947-8958
- José Pedro Iglesias, Amanda Nilsson, Carl Olsson: **expOSE: Accurate Initialization-Free Projective Factorization using Exponential Regularization.** 8959-8968
- Lin Huang, Chung-Ching Lin, Kevin Lin, Lin Liang, Lijuan Wang, Junsong Yuan, Zicheng Liu: **Neural Voting Field for Camera-Space 3D Hand Pose Estimation.** 8969-8978
- Axel Barroso-Laguna, Eric Brachmann, Victor Adrian Prisacariu, Gabriel J. Brostow, Daniyar Turmukhambetov: **Two-View Geometry Scoring Without Correspondences.** 8979-8989
- Petr Hruba, Viktor Korotynskiy, Timothy Duff , Luke Oeding, Marc Pollefeys, Tomás Pajdla, Viktor Larsson: **Four-view Geometry with Unknown Radial Distortion.** 8990-9000
- Jennifer J. Sun, Lili Karashchuk , Amil David, Serim Ryou, Sonia Fereidooni, John C. Tuthill, Aggelos K. Katsaggelos, Bingni W. Brunton, Georgia Gkioxari, Ann Kennedy, Yisong Yue, Pietro Perona: **BKinD-3D: Self-Supervised 3D Keypoint Discovery from Multi-View Videos.** 9001-9010
- Hyo-Jun Lee, Hanul Kim, Su-Min Choi, Seong-Gyun Jeong, Yeong Jun Koh: **BAAM: Monocular 3D pose and shape reconstruction with bi-contextual attention module and attention-guided modeling.** 9011-9020
- Stephen Tian, Yancheng Cai, Hong-Xing Yu, Sergey Zakharov, Katherine Liu, Adrien Gaidon, Yunzhu Li, Jiajun Wu: **Multi-Object Manipulation via Object-Centric Neural Scattering Functions.** 9021-9031
- Aleksei Bokhovkin, Angela Dai: **Neural Part Priors: Learning to Optimize Part-Based Object Completion in RGB-D Scans.** 9032-9042
- Yawar Siddiqui, Lorenzo Porzi, Samuel Rota Bulò, Norman Müller, Matthias Nießner, Angela Dai, Peter Kontschieder: **Panoptic Lifting for 3D Scene Understanding with Neural Fields.** 9043-9052
- Jamie Watson, Mohamed Sayed, Zawar Qureshi, Gabriel J. Brostow, Sara Vicente, Oisin Mac Aodha, Michael Firman: **Virtual Occlusions Through Implicit Depth.** 9053-9064
- Chao-Yuan Wu, Justin Johnson, Jitendra Malik, Christoph Feichtenhofer, Georgia Gkioxari: **Multiview Compressive Coding for 3D Reconstruction.** 9065-9075
- Felix Wimbauer, Nan Yang, Christian Rupprecht, Daniel Cremers: **Behind the Scenes: Density Fields for Single View Reconstruction.** 9076-9086
- Yiming Li, Zhiding Yu, Christopher B. Choy, Chaowei Xiao, José M. Álvarez, Sanja Fidler, Chen Feng, Anima Anandkumar: **VoxFormer: Sparse Voxel Transformer for Camera-Based 3D Semantic Scene Completion.** 9087-9098
- Obin Kwon, Jeongho Park, Songhwai Oh: **Renderable Neural Radiance Map for Visual Navigation.** 9099-9108
- Jiaying Lin , Xin Tan, Rynson W. H. Lau: **Learning to Detect Mirrors from Videos via Dual Correspondences.** 9109-9118
- Numair Khan, Eric Penner, Douglas Lanman, Lei Xiao: **Temporally Consistent Online Depth Estimation Using Point-Based Fusion.** 9119-9129
- Ruikang Xu, Mingde Yao, Zhiwei Xiong: **Zero-Shot Dual-Lens Super-Resolution.** 9130-9139



- Haozhe Si, Bin Zhao, Dong Wang, Yunpeng Gao, Mulin Chen, Zhigang Wang, Xuelong Li:
Fully Self-Supervised Depth Estimation from Defocus Clue. 9140-9149
- Xianggang Yu, Mutian Xu, Yidan Zhang, Haolin Liu , Chongjie Ye, Yushuang Wu, Zizheng Yan, Chenming Zhu, Zhangyang Xiong, Tianyou Liang, Guanying Chen , Shuguang Cui, Xiaoguang Han:
MVImgNet: A Large-scale Dataset of Multi-view Images. 9150-9161
- Ning Zhang, Yuyao Ye, Yang Zhao, Ronggang Wang:
Revisiting the Stack-Based Inverse Tone Mapping. 9162-9171
- Ruixuan Cong , Da Yang, Rongshan Chen, Sizhe Wang, Zhenglong Cui, Hao Sheng:
Combining Implicit-Explicit View Correlation for Light Field Semantic Segmentation. 9172-9181
- Mingtao Feng, Haoran Hou, Liang Zhang, Zijie Wu, Yulan Guo, Ajmal Mian :
3D Spatial Multimodal Knowledge Accumulation for Scene Graph Prediction in Point Cloud. 9182-9191
- Siddharth Somasundaram, Akshat Dave, Connor Henley , Ashok Veeraraghavan, Ramesh Raskar:
Role of Transients in Two-Bounce Non-Line-of-Sight Imaging. 9192-9201
- Yining Hong, Chunru Lin, Yilun Du, Zhenfang Chen, Joshua B. Tenenbaum, Chuang Gan:
3D Concept Learning and Reasoning from Multi-View Images. 9202-9212
- Dian Chen, Jie Li, Vitor Guizilini, Rares Ambrus, Adrien Gaidon:
Viewpoint Equivariance for Multi-View 3D Object Detection. 9213-9222
- Yuanhui Huang, Wenzhao Zheng, Yunpeng Zhang, Jie Zhou, Jiwen Lu:
Tri-Perspective View for Vision-Based 3D Semantic Occupancy Prediction. 9223-9232
- Wending Zhou, Xu Yan, Yinghong Liao, Yuankai Lin , Jin Huang, Gangming Zhao, Shuguang Cui, Zhen Li:
BEV@DC: Bird's-Eye View Assisted Training for Depth Completion. 9233-9242
- Yue Hu, Yifan Lu, Runsheng Xu, Weidi Xie, Siheng Chen, Yanfeng Wang:
Collaboration Helps Camera Overtake LiDAR in 3D Detection. 9243-9252
- Bo Zhang, Jiakang Yuan, Botian Shi, Tao Chen, Yikang Li, Yu Qiao:
Uni3D: A Unified Baseline for Multi-Dataset 3D Object Detection. 9253-9262
- Kemal Oksuz, Tom Joy, Puneet K. Dokania:
Towards Building Self-Aware Object Detectors via Reliable Uncertainty Quantification and Calibration. 9263-9274
- Akash Deep Singh, Yunhao Ba, Ankur Sarker, Howard Zhang, Achuta Kadambi, Stefano Soatto, Mani B. Srivastava, Alex Wong :
Depth Estimation from Camera Image and mmWave Radar Point Cloud. 9275-9285
- Wen Li, Shangshu Yu, Cheng Wang, Guosheng Hu, Siqi Shen, Chenglu Wen:
SGLoc: Scene Geometry Encoding for Outdoor LiDAR Localization. 9286-9295
- Benjin Zhu, Zhe Wang, Shaoshuai Shi, Hang Xu, Lanqing Hong, Hongsheng Li :
ConQueR: Query Contrast Voxel-DETR for 3D Object Detection. 9296-9305
- Chao Chen, Xinhao Liu, Yiming Li, Li Ding, Chen Feng:
DeepMapping2: Self-Supervised Large-Scale LiDAR Map Optimization. 9306-9316
- Lunjun Zhang, Anqi Joyce Yang, Yuwen Xiong, Sergio Casas, Bin Yang, Mengye Ren, Raquel Urtasun:
Towards Unsupervised Object Detection from LiDAR Point Clouds. 9317-9328
- Yingwei Li, Charles R. Qi, Yin Zhou, Chenxi Liu, Dragomir Anguelov:
MoDAR: Using Motion Forecasting for 3D Object Detection in Point Cloud Sequences. 9329-9339
- Fangqiang Ding, Andras Palfy , Dariu M. Gavrila, Chris XiaoXuan Lu:
Hidden Gems: 4D Radar Scene Flow Learning Using Cross-Modal Supervision. 9340-9349
- Kwonyoung Ryu, Soonmin Hwang , Jaesik Park:
Instant Domain Augmentation for LiDAR Semantic Segmentation. 9350-9360
- Li Li, Hubert P. H. Shum , Toby P. Breckon:
Less is More: Reducing Task and Model Complexity for 3D Point Cloud Semantic Segmentation. 9361-9371
- Jiahui Liu, Chirui Chang, Jianhui Liu, Xiaoyang Wu, Lan Ma, Xiaojuan Qi:
MarS3D: A Plug-and-Play Motion-Aware Model for Semantic Segmentation on Multi-Scan 3D Point Clouds. 9372-9381
- Aoran Xiao, Jiaxing Huang, Weihao Xuan, Ruijie Ren, Kangcheng Liu, Dayan Guan, Abdulmotaleb El-Saddik, Shijian Lu, Eric P. Xing:
3D Semantic Segmentation in the Wild: Learning Generalized Models for Adverse-Condition Point Clouds. 9382-9392
- Luigi Riz , Cristiano Saltori, Elisa Ricci, Fabio Poiesi:
Novel Class Discovery for 3D Point Cloud Semantic Segmentation. 9393-9402
- Honghui Yang, Tong He, Jiaheng Liu, Hua Chen, Boxi Wu, Binbin Lin, Xiaofei He, Wanli Ouyang:
GD-MAE: Generative Decoder for MAE Pre-Training on LiDAR Point Clouds. 9403-9414
- Xiaoyang Wu, Xin Wen , Xihui Liu, Hengshuang Zhao:
Masked Scene Contrast: A Scalable Framework for Unsupervised 3D Representation Learning. 9415-9424



- Jianan Li, Qiulei Dong:
Open-set Semantic Segmentation for Point Clouds via Adversarial Prototype Framework. 9425-9434
- Sangmin Hong, Mohsen Yavartanoo, Reyhaneh Neshatavar, Kyoung Mu Lee:
ACL-SPC: Adaptive Closed-Loop System for Self-Supervised Point Cloud Completion. 9435-9444
- Lemeng Wu, Dilin Wang, Chengyue Gong, Xingchao Liu, Yunyang Xiong, Rakesh Ranjan, Raghuraman Krishnamoorthi, Vikas Chandra, Qiang Liu:
Fast Point Cloud Generation with Straight Flows. 9445-9454
- Xin Deng, Wenyu Zhang , Qing Ding, Ximeng Zhang :
PointVector: A Vector Representation In Point Cloud Analysis. 9455-9465
- Shanshan Li, Pan Gao, Xiaoyang Tan, Mingqiang Wei:
ProxyFormer: Proxy Alignment Assisted Point Cloud Completion with Missing Part Sensitive Transformer. 9466-9475
- Kangcheng Liu, Aoran Xiao, Xiaoqin Zhang, Shijian Lu, Ling Shao:
FAC: 3D Representation Learning via Foreground Aware Feature Contrast. 9476-9485
- Hang Du, Xuejun Yan, Jingjing Wang, Di Xie, Shiliang Pu:
Rethinking the Approximation Error in 3D Surface Fitting for Point Cloud Normal Estimation. 9486-9495
- Jinghuai Zhang, Jinyuan Jia, Hongbin Liu, Neil Zhenqiang Gong:
PointCert: Point Cloud Classification with Deterministic Certified Robustness Guarantees. 9496-9505
- Haiping Wang , Yuan Liu, Zhen Dong, Yulan Guo, Yu-Shen Liu, Wenping Wang, Bisheng Yang:
Robust Multiview Point Cloud Registration with Reliable Pose Graph Initialization and History Reweighting. 9506-9515
- Jiawen Zhu, Simiao Lai , Xin Chen, Dong Wang, Huchuan Lu:
Visual Prompt Multi-Modal Tracking. 9516-9526
- Xin Liu, Jufeng Yang:
Progressive Neighbor Consistency Mining for Correspondence Pruning. 9527-9537
- Yuting He, Guanyu Yang, Rongjun Ge, Yang Chen, Jean-Louis Coatrieux, Boyu Wang, Shuo Li :
Geometric Visual Similarity Learning in 3D Medical Image Self-Supervised Pre-training. 9538-9547
- Zesen Wu, Mang Ye :
Unsupervised Visible-Infrared Person Re-Identification via Progressive Graph Matching and Alternate Learning. 9548-9558
- Tianyu Chang, Xun Yang, Tianzhu Zhang, Meng Wang:
Domain Generalized Stereo Matching via Hierarchical Visual Transformation. 9559-9568
- Hanyu Zhou, Yi Chang, Wending Yan, Luxin Yan:
Unsupervised Cumulative Domain Adaptation for Foggy Scene Optical Flow. 9569-9578
- Weicai Ye, Xinyue Lan, Shuo Chen, Yuhang Ming , Xingyuan Yu, Hujun Bao, Zhaopeng Cui, Guofeng Zhang:
PVO: Panoptic Visual Odometry. 9579-9589
- Cong Pan, Yonghao He, Junran Peng, Qian Zhang, Wei Sui, Zhaoxiang Zhang:
BAEFormer: Bi-Directional and Early Interaction Transformers for Bird's Eye View Semantic Segmentation. 9590-9599
- Xiaofeng Wang, Zheng Zhu, Yunpeng Zhang, Guan Huang, Yun Ye, Wenbo Xu, Ziwei Chen, Xingang Wang:
Are We Ready for Vision-Centric Driving Streaming Perception? The ASAP Benchmark. 9600-9610
- Xiwen Liang, Minzhe Niu, Jianhua Han, Hang Xu, Chunjing Xu, Xiaodan Liang:
Visual Exemplar Driven Task-Prompting for Unified Perception in Autonomous Driving. 9611-9621
- Simon Suo, Kelvin Wong, Justin Xu, James Tu, Alexander Cui, Sergio Casas, Raquel Urtasun:
MIXSIM: A Hierarchical Framework for Mixed Reality Traffic Simulation. 9622-9631
- Yi Xu, Armin Bazarjani, Hyung-Gun Chi, Chiho Choi, Yun Fu:
Uncovering the Missing Pattern: Unified Framework Towards Trajectory Imputation and Prediction. 9632-9643
- Chiyu Max Jiang, Andre Cornman, Cheolho Park, Benjamin Sapp, Yin Zhou, Dragomir Anguelov:
MotionDiffuser: Controllable Multi-Agent Motion Prediction Using Diffusion. 9644-9653
- Sammy Joe Christen, Wei Yang, Claudia Pérez-D'Arpino, Otmar Hilliges, Dieter Fox, Yu-Wei Chao:
Learning Human-to-Robot Handovers from Point Clouds. 9654-9664
- Matt Deitke, Rose Hendrix, Ali Farhadi, Kiana Ehsani, Aniruddha Kembhavi:
Phone2Proc: Bringing Robust Robots into Our Chaotic World. 9665-9675
- Alessandro Ruzzi, Xiangwei Shi, Xi Wang, Gengyan Li, Shalini De Mello, Hyung Jin Chang, Xucong Zhang, Otmar Hilliges:
GazeNeRF: 3D-Aware Gaze Redirection with Neural Radiance Fields. 9676-9685
- Jinkun Cao, Jiangmiao Pang, Xinshuo Weng, Rawal Khirodkar, Kris Kitani:
Observation-Centric SORT: Rethinking SORT for Robust Multi-Object Tracking. 9686-9696
- Xing Wei, Yifan Bai, Yongchao Zheng, Dahu Shi, Yihong Gong:
Autoregressive Visual Tracking. 9697-9706
- Chao Fan, Junhao Liang , Chuanfu Shen, Saihui Hou, Yongzhen Huang, Shiqi Yu:
OpenGait: Revisiting Gait Recognition Toward Better Practicality. 9707-9716



- Yuanyuan Liu, Wenbin Wang, Yibing Zhan, Shaoze Feng, Kejun Liu, Zhe Chen : **Pose-disentangled Contrastive Learning for Self-supervised Facial Representation.** 9717-9728
- Weizhi Zhong, Chaowei Fang, Yinqi Cai, Pengxu Wei, Gangming Zhao, Liang Lin, Guanbin Li: **Identity-Preserving Talking Face Generation with Landmark and Appearance Priors.** 9729-9738
- Kartik Narayan, Harsh Agarwal, Kartik Thakral, Surbhi Mittal , Mayank Vatsa, Richa Singh: **DF-Platter: Multi-Face Heterogeneous Deepfake Dataset.** 9739-9748
- Kun Su, Kaizhi Qian, Eli Shlizerman, Antonio Torralba, Chuang Gan: **Physics-Driven Diffusion Models for Impact Sound Synthesis from Videos.** 9749-9759
- Rishabh Dabral, Muhammad Hamza Mughal, Vladislav Golyanik, Christian Theobalt: **MoFusion: A Framework for Denoising-Diffusion-Based Motion Synthesis.** 9760-9770
- Urbano Miguel Nunes, Ryad Benosman, Sio-Hoi Ieng: **Adaptive Global Decay Process for Event Cameras.** 9771-9780
- Jiqing Zhang, Yuanchen Wang, Wenxi Liu, Meng Li, Jinpeng Bai, Baocai Yin, Xin Yang: **Frame-Event Alignment and Fusion Network for High Frame Rate Tracking.** 9781-9790
- Sangjin Lee, Hyeongmin Lee, Chajin Shin, Hanbin Son, Sangyoun Lee: **Exploring Discontinuity for Video Frame Interpolation.** 9791-9800
- Zhen Li, Zuo-Liang Zhu, Linghao Han, Qibin Hou, Chun-Le Guo, Ming-Ming Cheng: **AMT: All-Pairs Multi-Field Transforms for Efficient Frame Interpolation.** 9801-9810
- Markus Plack, Matthias B. Hullin, Karlis Martins Briedis, Markus Gross , Abdelaziz Djelouah, Christopher Schroers: **Frame Interpolation Transformer and Uncertainty Guidance.** 9811-9821
- Dasong Li, Xiaoyu Shi, Yi Zhang, Ka Chun Cheung, Simon See, Xiaogang Wang, Hongwei Qin, Hongsheng Li: **A Simple Baseline for Video Restoration with Grouped Spatial-Temporal Shift.** 9822-9832
- Si-Yuan Cao, Runmin Zhang , Lun Luo, Beinan Yu, Zehua Sheng, Junwei Li, Hui-Liang Shen: **Recurrent Homography Estimation Using Homography-Guided Image Warping and Focus Transformer.** 9833-9842
- Bang-Dang Pham , Phong Tran, Anh Tran, Cuong Pham, Rang Nguyen, Minh Hoai: **HyperCUT: Video Sequence from a Single Blurry Image using Unsupervised Ordering.** 9843-9852
- Lingke Kong, X. Sharon Qi, Qijin Shen, Jiacheng Wang, Jingyi Zhang, Yanle Hu, Qichao Zhou: **Indescribable Multi-Modal Spatial Evaluator.** 9853-9862
- Yash Sanghvi, Zhiyuan Mao, Stanley H. Chan: **Structured Kernel Estimation for Photon-Limited Deconvolution.** 9863-9872
- Zhuoxiao Li, Haiyang Jiang, Mingdeng Cao, Yingqiang Zheng: **Polarized Color Image Denoising.** 9873-9882
- Xiaole Tang , Xile Zhao, Jun Liu, Jianli Wang, Yuchun Miao, Tieyong Zeng: **Uncertainty-Aware Unsupervised Image Deblurring with Deep Residual Prior.** 9883-9892
- Xiaogang Xu, Ruixing Wang, Jiangbo Lu: **Low-Light Image Enhancement via Structure Modeling and Guidance.** 9893-9903
- Jie Huang, Feng Zhao, Man Zhou, Jie Xiao, Naishan Zheng, Kaiwen Zheng, Zhiwei Xiong: **Learning Sample Relationship for Exposure Correction.** 9904-9913
- Junyi Li, Zhilu Zhang, Xiaoyu Liu, Chaoyu Feng, Xiaotao Wang, Lei Lei, Wangmeng Zuo: **Spatially Adaptive Self-Supervised Learning for Real-World Image Denoising.** 9914-9924
- Jie Zhang, Yongshan Zhang, Yicong Zhou: **Quantum-Inspired Spectral-Spatial Pyramid Network for Hyperspectral Image Classification.** 9925-9934
- Ben Fei, Zhaoyang Lyu, Liang Pan, Junzhe Zhang, Weidong Yang , Tianyue Luo, Bo Zhang , Bo Dai: **Generative Diffusion Prior for Unified Image Restoration and Enhancement.** 9935-9946
- Xinran Qin, Yuhui Quan, Tongyao Pang, Hui Ji: **Ground-Truth Free Meta-Learning for Deep Compressive Sampling.** 9947-9956
- Jacky Chen Long Chai, Tiong-Sik Ng, Cheng-Yaw Low, Jaewoo Park, Andrew Beng Jin Teoh: **Recognizability Embedding Enhancement for Very Low-Resolution Face Recognition and Quality Estimation.** 9957-9967
- Nicolas Chahine, Ana-Stefania Calarasanu, Davide Garcia-Civiero, Théo Cayla, Sira Ferradans, Jean Ponce: **An Image Quality Assessment Dataset for Portraits.** 9968-9978
- Wenyang Liu, Yi Wang, Kim-Hui Yap, Lap-Pui Chau: **Bitstream-Corrupted JPEG Images are Restorable: Two-stage Compensation and Alignment Framework for Image Restoration.** 9979-9988
- Simon Grosche, Andy Regensky, Jürgen Seiler, André Kaup: **Image Super-Resolution Using T-Tetromino Pixels.** 9989-9998
- Cristina Nader Vasconcelos, A. Cengiz Öztireli, Mark J. Matthews , Milad Hashemi, Kevin Swersky, Andrea Tagliasacchi: **CUF: Continuous Upsampling Filters.** 9999-10008



- Gaochao Song, Qian Sun, Luo Zhang, Ran Su, Jianfeng Shi, Ying He:
OPE-SR: Orthogonal Position Encoding for Designing a Parameter-free Upsampling Module in Arbitrary-scale Image Super-Resolution. 10009-10020
- Sicheng Gao, Xuhui Liu, Bohan Zeng, Sheng Xu, Yanjing Li, Xiaoyan Luo, Jianzhuang Liu, Xiantong Zhen, Baochang Zhang:
Implicit Diffusion Models for Continuous Super-Resolution. 10021-10030
- Yi Wang, Ruili Wang, Xin Fan, Tianzhu Wang, Xiangjian He :
Pixels, Regions, and Objects: Multiple Enhancement for Salient Object Detection. 10031-10040
- Junjie Ke, Keren Ye, Jiahui Yu, Yonghui Wu, Peyman Milanfar, Feng Yang:
VILA: Learning Image Aesthetics from User Comments with Vision-Language Pretraining. 10041-10051
- Chao Wang, Li Niu, Bo Zhang, Liqing Zhang:
Image Cropping with Spatial-aware Feature and Rank Consistency. 10052-10061
- Byeonghyun Pak, Jaewon Lee, Kyong Hwan Jin:
B-Spline Texture Coefficients Estimator for Screen Content Image Super-Resolution. 10062-10071
- Hongyu Liu, Yibing Song, Qifeng Chen:
Delving StyleGAN Inversion for Image Editing: A Foundation Latent Space Viewpoint. 10072-10082
- Wenju Xu, Chengjiang Long, Yongwei Nie:
Learning Dynamic Style Kernels for Artistic Style Transfer. 10083-10092
- Defu Cao , Zhaojun Wang, Jose Echevarria, Yan Liu:
SVGformer: Representation Learning for Continuous Vector Graphics using Transformers. 10093-10102
- Xiaoming Li, Wangmeng Zuo, Chen Change Loy:
Learning Generative Structure Prior for Blind Text Image Super-resolution. 10103-10113
- Chenchen Xu, Min Zhou, Tiezheng Ge, Yuning Jiang, Weiwei Xu:
Unsupervised Domain Adaptation with Pixel-Level Discriminator for Image-Aware Layout Generation. 10114-10123
- Minguk Kang, Jun-Yan Zhu, Richard Zhang, Jaesik Park, Eli Shechtman, Sylvain Paris, Taesung Park:
Scaling up GANs for Text-to-Image Synthesis. 10124-10134
- Zhida Feng, Zhenyu Zhang, Xintong Yu, Yewei Fang, Lanxin Li, Xuyi Chen, Yuxiang Lu, Jiaxiang Liu, Weichong Yin, Shikun Feng, Yu Sun, Li Chen, Hao Tian, Hua Wu, Haifeng Wang:
ERNIE-ViLG 2.0: Improving Text-to-Image Diffusion Model with Knowledge-Enhanced Mixture-of-Denoising Experts. 10135-10145
- Yuxin Zhang, Nisha Huang, Fan Tang, Haibin Huang, Chongyang Ma, Weiming Dong, Changsheng Xu:
Inversion-based Style Transfer with Diffusion Models. 10146-10156
- Yufan Zhou, Bingchen Liu, Yizhe Zhu, Xiao Yang, Changyou Chen, Jinhui Xu:
Shifted Diffusion for Text-to-image Generation. 10157-10166
- Naoto Inoue, Kotaro Kikuchi, Edgar Simo-Serra, Mayu Otani, Kota Yamaguchi:
LayoutDM: Discrete Diffusion Model for Controllable Layout Generation. 10167-10176
- Shaoan Xie, Yanwu Xu, Mingming Gong, Kun Zhang:
Unpaired Image-to-Image Translation with Shortest Path Regularization. 10177-10187
- Qinsheng Zhang, Jiaming Song, Xun Huang, Yongxin Chen, Ming-Yu Liu:
DiffCollage: Parallel Generation of Large Content with Diffusion Models. 10188-10198
- Hao Phung, Quan Dao, Anh Tran:
Wavelet Diffusion Models are fast and scalable Image Generators. 10199-10208
- (Withdrawn) VideoFusion: Decomposed Diffusion Models for High-Quality Video Generation. 10209-10218
- Ludan Ruan, Yiyang Ma, Huan Yang, Huiguo He, Bei Liu, Jianlong Fu, Nicholas Jing Yuan, Qin Jin, Baining Guo:
MM-Diffusion: Learning Multi-Modal Diffusion Models for Joint Audio and Video Generation. 10219-10228
- Chung-Ching Lin, Jiang Wang, Kun Luo, Kevin Lin, Linjie Li, Lijuan Wang, Zicheng Liu:
Adaptive Human Matting for Dynamic Videos. 10229-10238
- Xi Zhang , Xiaolin Wu:
LVQAC: Lattice Vector Quantization Coupled with Spatially Adaptive Companding for Efficient Learned Image Compression. 10239-10248
- David Alexandre, Hsueh-Ming Hang, Wen-Hsiao Peng:
Hierarchical B-Frame Video Coding Using Two-Layer CANF Without Motion Coding. 10249-10258
- Gen Li, Jie Ji, Minghai Qin, Wei Niu, Bin Ren, Fatemeh Afghah, Linke Guo, Xiaolong Ma:
Towards High-Quality and Efficient Video Super-Resolution via Spatial-Temporal Data Overfitting. 10259-10269
- Hao Chen, Matthew Gwilliam, Ser-Nam Lim, Abhinav Shrivastava:
HNeRV: A Hybrid Neural Representation for Videos. 10270-10279
- Zhemin Li, Hongxia Wang, Deyu Meng:
Regularize implicit neural representation by itself. 10280-10288
- Sanghyeon Kim, Eunbyung Park:
SMPConv: Self-Moving Point Representations for Continuous Convolution. 10289-10299



- Xiang-Li Li, Meng-Hao Guo, Tai-Jiang Mu , Ralph R. Martin, Shi-Min Hu:
Long Range Pooling for 3D Large-Scale Scene Understanding. 10300-10311
- Seokeon Choi, Debasmit Das, Sungjae Choi, Seunghan Yang, Hyunsin Park, Sungrack Yun:
Progressive Random Convolutions for Single Domain Generalization. 10312-10322
- Lei Zhu , Xinjiang Wang, Zhanghan Ke, Wayne Zhang, Rynson W. H. Lau:
BiFormer: Vision Transformer with Bi-Level Routing Attention. 10323-10333
- Sifan Long , Zhen Zhao, Jimin Pi, Shengsheng Wang, Jingdong Wang:
Beyond Attentive Tokens: Incorporating Token Importance and Diversity for Efficient Vision Transformers. 10334-10343
- Pengyu Li:
BioNet: A Biologically-Inspired Network for Face Recognition. 10344-10354
- Jingda Du , Siqi Liu, Bochao Zhang, Pong C. Yuen:
Dual-bridging with Adversarial Noise Generation for Domain Adaptive rPPG Estimation. 10355-10364
- Zhenda Xie, Zheng Zhang, Yue Cao, Yutong Lin, Yixuan Wei, Qi Dai, Han Hu:
On Data Scaling in Masked Image Modeling. 10365-10374
- Haochen Wang, Kaiyou Song, Junsong Fan, Yuxi Wang, Jin Xie, Zhaoxiang Zhang:
Hard Patches Mining for Masked Image Modeling. 10375-10385
- Zhanzhou Feng, Shiliang Zhang:
Evolved Part Masking for Self-Supervised Learning. 10386-10395
- Or Streicher, Ido Cohen, Guy Gilboa:
BASiS: Batch Aligned Spectral Embedding Space. 10396-10405
- Rohit Girdhar, Alaaddin El-Nouby, Mannat Singh, Kalyan Vasudev Alwala, Armand Joulin, Ishan Misra:
OmniMAE: Single Model Masked Pretraining on Images and Videos. 10406-10417
- Michail Tarasiou, Erik Chavez, Stefanos Zafeiriou:
ViTs for SITS: Vision Transformers for Satellite Image Time Series. 10418-10428
- Bashirul Azam Biswas, Qiang Ji:
Probabilistic Debiasing of Scene Graphs. 10429-10438
- Chenyang Lei, Xuanchi Ren , Zhaoxiang Zhang, Qifeng Chen:
Blind Video Deflickering by Neural Filtering with a Flawed Atlas. 10439-10448
- Lihao Liu, Jean Prost, Lei Zhu, Nicolas Papadakis, Pietro Liò, Carola-Bibiane Schönlieb, Angelica I. Avilés-Rivero:
SCOTCH and SODA: A Transformer Video Shadow Detection Framework. 10449-10458
- Lijun Yu, Yong Cheng, Kihyuk Sohn, José Lezama, Han Zhang, Huiwen Chang, Alexander G. Hauptmann, Ming-Hsuan Yang, Yuan Hao, Irfan Essa, Lu Jiang:
MAGVIT: Masked Generative Video Transformer. 10459-10469
- Aakanksha, A. N. Rajagopalan:
Improving Robustness of Semantic Segmentation to Motion-Blur Using Class-Centric Augmentation. 10470-10479
- Roy Miles, Mehmet Kerim Yucel , Bruno Manganelli, Albert Saà-Garriga:
MobileVOS: Real-Time Video Object Segmentation Contrastive Learning meets Knowledge Distillation. 10480-10490
- Chao Feng, Ziyang Chen, Andrew Owens:
Self-Supervised Video Forensics by Audio-Visual Anomaly Detection. 10491-10503
- Yitian Zhang, Yue Bai, Chang Liu, Huan Wang, Sheng Li, Yun Fu:
Frame Flexible Network. 10504-10513
- Lin Geng Foo, Jia Gong, Zhipeng Fan , Jun Liu:
System-Status-Aware Adaptive Network for Online Streaming Video Understanding. 10514-10523
- Minghan Li , Shuai Li, Wangmeng Xiang, Lei Zhang:
MDQE: Mining Discriminative Query Embeddings to Segment Occluded Instances on Challenging Videos. 10524-10533
- Shao-Yuan Lo, Poojan Oza, Sumanth Chennupati, Alejandro Galindo, Vishal M. Patel:
Spatio-Temporal Pixel-Level Contrastive Learning-based Source-Free Domain Adaptation for Video Semantic Segmentation. 10534-10543
- Lingting Zhu, Xian Liu, Xuanyu Liu, Rui Qian, Ziwei Liu, Lequan Yu:
Taming Diffusion Models for Audio-Driven Co-Speech Gesture Generation. 10544-10553
- Sagnik Majumder, Hao Jiang, Pierre Moulon, Ethan Henderson, Paul Calamia, Kristen Grauman, Vamsi Krishna Ithapu:
Chat2Map: Efficient Scene Mapping from Multi-Ego Conversations. 10554-10564
- Shentong Mo, Yapeng Tian:
Audio-Visual Grouping Network for Sound Localization from Mixtures. 10565-10574
- Reuben Tan, Arijit Ray, Andrea Burns, Bryan A. Plummer, Justin Salamon, Oriol Nieto, Bryan Russell, Kate Saenko:
Language-Guided Audio-Visual Source Separation via Trimodal Consistency. 10575-10584
- Xuyang Shen, Dong Li, Jinxing Zhou, Zhen Qin, Bowen He, Xiaodong Han, Aixuan Li, Yuchao Dai, Lingpeng Kong, Meng Wang, Yu Qiao, Yiran Zhong:
https://dblp.org/db/conf/cvpr/cvpr2023.html

**Fine-grained Audible Video Description.** 10585-10596

Xinghan Wang, Xin Xu, Yadong Mu:

Neural Koopman Pooling: Control-Inspired Temporal Dynamics Encoding for Skeleton-Based Action Recognition. 10597-10607

Huanyu Zhou, Qingjie Liu, Yunhong Wang:

Learning Discriminative Representations for Skeleton Based Action Recognition. 10608-10617

Eadom Dessaalene, Michael Maynord, Cornelia Fermüller, Yiannis Aloimonos:

Therbligs in Action: Video Understanding through Motion Primitives. 10618-10626

Mingjun Zhao, Yakun Yu, Xiaoli Wang, Lei Yang, Di Niu:

Search-Map-Search: A Frame Selection Paradigm for Action Recognition. 10627-10636

Chen Zhao, Shuming Liu, Karttikeya Mangalam, Bernard Ghanem

Re²TAL: Rewiring Pretrained Video Backbones for Reversible Temporal Action Localization. 10637-10647

Guozhang Li , De Cheng, Xinpeng Ding, Nannan Wang, Xiaoyu Wang, Xinbo Gao:

Boosting Weakly-Supervised Temporal Action Localization with Text Information. 10648-10657

Zhenghua Peng, Yu Luo, Tianshui Chen, Keke Xu, Shuangping Huang:

Perception and Semantic Aware Regularization for Sequential Confidence Calibration. 10658-10668

Haoqian Wu, Keyu Chen, Haozhe Liu, Mingchen Zhuge, Bing Li, Ruizhi Qiao, Xiujun Shu, Bei Gan, Liangsheng Xu, Bo Ren,

Mengmeng Xu, Wentian Zhang, Raghavendra Ramachandra, Chia-Wen Lin, Bernard Ghanem

NewsNet: A Novel Dataset for Hierarchical Temporal Segmentation. 10669-10680

Tsui-Jui Fu, Licheng Yu, Ning Zhang, Cheng-Yang Fu, Jong-Chyi Su, William Yang Wang, Sean Bell:

Tell Me What Happened: Unifying Text-guided Video Completion via Multimodal Masked Video Generation. 10681-10692

Camilo Luciano Fosco, SouYoung Jin, Emilie Josephs, Aude Oliva:

Leveraging Temporal Context in Low Representational Power Regimes. 10693-10703

Wenhai Wu , Haipeng Luo, Bo Fang, Jingdong Wang, Wanli Ouyang:

Cap4Video: What Can Auxiliary Captions Do for Text-Video Retrieval? 10704-10713

Antoine Yang, Arsha Nagrani, Paul Hongsuck Seo, Antoine Miech, Jordi Pont-Tuset, Ivan Laptev, Josef Sivic, Cordelia Schmid:

Vid2Seq: Large-Scale Pretraining of a Visual Language Model for Dense Video Captioning. 10714-10726

Honglu Zhou, Roberto Martín-Martín, Mubbasis Kapadia, Silvio Savarese, Juan Carlos Niebles:

Procedure-Aware Pretraining for Instructional Video Understanding. 10727-10738

Feng Cheng, Xizi Wang, Jie Lei, David J. Crandall, Mohit Bansal, Gedas Bertasius:

VindLU: A Recipe for Effective Video-and-Language Pretraining. 10739-10750

Théo Dumont, Juan Segundo Hevia, Camilo Luciano Fosco:

Modular Memorability: Tiered Representations for Video Memorability Prediction. 10751-10760

Feiyu Chen, Jie Shao, Shuyuan Zhu, Heng Tao Shen:

Multivariate, Multi-Frequency and Multimodal: Rethinking Graph Neural Networks for Emotion Recognition in Conversation. 10761-10770

Leming Guo , Wanli Xue, Qing Guo, Bo Liu, Kaihua Zhang, Tiantian Yuan, Shengyong Chen:

Distilling Cross-Temporal Contexts for Continuous Sign Language Recognition. 10771-10780

Shengkun Tang, Yaqing Wang, Zhenglun Kong, Tianchi Zhang, Yao Li, Caiwen Ding, Yanzhi Wang, Yi Liang, Dongkuan Xu:

You Need Multiple Exiting: Dynamic Early Exiting for Accelerating Unified Vision Language Model. 10781-10791

Sixian Zhang, Xinhang Song, Weijie Li, Yubing Bai, Xinyao Yu, Shuqiang Jiang:

Layout-based Causal Inference for Object Navigation. 10792-10802

Jialu Li, Mohit Bansal:

Improving Vision-and-Language Navigation by Generating Future-View Image Semantics. 10803-10812

Aishwarya Kamath, Peter Anderson, Su Wang, Jing Yu Koh, Alexander Ku, Austin Waters, Yinfei Yang, Jason Baldridge, Zarana Parekh:

A New Path: Scaling Vision-and-Language Navigation with Synthetic Instructions and Imitation Learning. 10813-10823

Duc Minh Vo, Quoc-An Luong, Akihiro Sugimoto, Hideki Nakayama:

A-CAP: Anticipation Captioning with Commonsense Knowledge. 10824-10833

Anoop Cherian, Kuan-Chuan Peng, Suhas Lohit, Kevin A. Smith, Joshua B. Tenenbaum:

Are Deep Neural Networks SMARTer Than Second Graders? 10834-10844

Youngjae Yu, Jiwan Chung, Heeseung Yun, Jack Hessel, Jae Sung Park, Ximing Lu, Rowan Zellers, Prithviraj Ammanabrolu, Ronan Le Bras, Gunhee Kim, Yejin Choi:

Fusing Pre-Trained Language Models with Multimodal Prompts through Reinforcement Learning. 10845-10856

Wei Su, Peihan Miao, Huanzhang Dou, Gaoang Wang, Liang Qiao, Zheyang Li, Xi Li:

Language Adaptive Weight Generation for Multi-Task Visual Grounding. 10857-10866

Jiaxian Guo, Junnan Li, Dongxu Li, Anthony Meng Huat Tiong, Boyang Li, Dacheng Tao, Steven C. H. Hoi:

From Images to Textual Prompts: Zero-shot Visual Question Answering with Frozen Large Language Models.



10867-10877

Qidong Huang, Xiaoyi Dong, Dongdong Chen, Weiming Zhang, Feifei Wang, Gang Hua, Nenghai Yu:
Diversity-Aware Meta Visual Prompting. 10878-10887

Yajing Liu, Yuning Lu, Hao Liu, Yaozu An, Zhuoran Xu, Zhuokun Yao, Baofeng Zhang, Zhiwei Xiong, Chenguang Gui:
Hierarchical Prompt Learning for Multi-Task Learning. 10888-10898

Tao Yu, Zhihe Lu, Xin Jin, Zhibo Chen, Xinchao Wang:
Task Residual for Tuning Vision-Language Models. 10899-10909

Zixian Ma, Jerry Hong, Mustafa Omer Gul, Mona Gandhi, Irena Gao, Ranjay Krishna:
@ CREPE: Can Vision-Language Foundation Models Reason Compositionally? 10910-10921

Gen Li, Varun Jampani, Deqing Sun, Laura Sevilla-Lara:
LOCATE: Localize and Transfer Object Parts for Weakly Supervised Affordance Grounding. 10922-10931

Vikram V. Ramaswamy , Sunnie S. Y. Kim, Ruth Fong, Olga Russakovsky:
Overlooked Factors in Concept-Based Explanations: Dataset Choice, Concept Learnability, and Human Capability. 10932-10941

Siwon Kim, Jinoh Oh, Sungjin Lee, Seunghak Yu, Jaeyoung Do, Tara Taghavi:
Grounding Counterfactual Explanation of Image Classifiers to Textual Concept Space. 10942-10950

Da Yin, Feng Gao, Govind Thattai, Michael Johnston, Kai-Wei Chang:
GIVL: Improving Geographical Inclusivity of Vision-Language Models with Pre-Training Methods. 10951-10961

Bowen Wang , Liangzhi Li, Yuta Nakashima, Hajime Nagahara:
Learning Bottleneck Concepts in Image Classification. 10962-10971

Pinaki Nath Chowdhury, Ayan Kumar Bhunia, Aneeshan Sain , Subhadeep Koley, Tao Xiang, Yi-Zhe Song :
SceneTrilogy: On Human Scene-Sketch and its Complementarity with Photo and Text. 10972-10983

Zhao Jin, Munawar Hayat, Yuwei Yang, Yulan Guo, Yinjie Lei:
Context-aware Alignment and Mutual Masking for 3D-Language Pre-training. 10984-10994

Xiaoyi Dong, Jianmin Bao, Yinglin Zheng, Ting Zhang, Dongdong Chen, Hao Yang, Ming Zeng, Weiming Zhang, Lu Yuan, Dong Chen, Fang Wen, Nenghai Yu:
MaskCLIP: Masked Self-Distillation Advances Contrastive Language-Image Pretraining. 10995-11005

Michael Tschannen, Basil Mustafa, Neil Houlsby:
CLIPPO: Image-and-Language Understanding from Pixels Only. 11006-11017

Yuxin Chen, Zongyang Ma, Ziqi Zhang, Zhongang Qi, Chunfeng Yuan, Ying Shan, Bing Li, Weiming Hu, Xiaohu Qie, Jianping Wu:
ViLEM: Visual-Language Error Modeling for Image-Text Retrieval. 11018-11027

Jinghao Zhou, Li Dong, Zhe Gan, Lijuan Wang, Furu Wei:
Non-Contrastive Learning Meets Language-Image Pre-Training. 11028-11038

Chia-Wen Kuo, Zsolt Kira:
HAAV: Hierarchical Aggregation of Augmented Views for Image Captioning. 11039-11049

Yang Jiao, Yan Gao, Jingjing Meng, Jin Shang, Yi Sun:
Learning Attribute and Class-Specific Representation Duet for Fine-Grained Fashion Analysis. 11050-11059

Yang Jin, Yongzhi Li, Zehuan Yuan, Yadong Mu:
Learning Instance-Level Representation for Large-Scale Multi-Modal Pretraining in E-Commerce. 11060-11069

Dmytro Kotovenko, Pingchuan Ma, Timo Milbich, Björn Ommer:
Cross-Image-Attention for Conditional Embeddings in Deep Metric Learning. 11070-11081

Hui Wu, Min Wang, Wengang Zhou, Zhenbo Lu, Houqiang Li:
Asymmetric Feature Fusion for Image Retrieval. 11082-11092

Yunhao Ge, Jie Ren, Andrew Gallagher, Yuxiao Wang, Ming-Hsuan Yang, Hartwig Adam, Laurent Itti, Balaji Lakshminarayanan, Jiaping Zhao:
Improving Zero-shot Generalization and Robustness of Multi-Modal Models. 11093-11101

Zhongzhi Yu, Shang Wu, Yonggan Fu, Shunyao Zhang, Yingyan Celine Lin:
Hint-Aug: Drawing Hints from Foundation Vision Transformers towards Boosted Few-shot Parameter-Efficient Tuning. 11102-11112

Benjamin Ramtoula, Matthew Gadd, Paul Newman, Daniele De Martini:
Visual DNA: Representing and Comparing Images Using Distributions of Neuron Activations. 11113-11123

Sijin Chen, Hongyuan Zhu, Xin Chen , Yinjie Lei, Gang Yu, Tao Chen:
End-to-End 3D Dense Captioning with Vote2Cap-DETR. 11124-11133

Yongshuai Huang, Ning Lu, Dapeng Chen, Yibo Li, Zecheng Xie, Shenggao Zhu, Liangcai Gao, Wei Peng:
Improving Table Structure Recognition with Visual-Alignment Sequential Coordinate Modeling. 11134-11143

Dahun Kim, Anelia Angelova, Weicheng Kuo:
Region-Aware Pretraining for Open-Vocabulary Object Detection with Vision Transformers. 11144-11154

Zhangxuan Gu, Zhuoer Xu, Haoxing Chen, Jun Lan, Changhua Meng, Weiqiang Wang:
Mobile User Interface Element Detection Via Adaptively Prompt Tuning. 11155-11164



- Junbum Cha, Jonghwan Mun, Byungseok Roh:
Learning to Generate Text-Grounded Mask for Open-World Semantic Segmentation from Only Image-Text Pairs. 11165-11174
- Ziqin Zhou, Yinjie Lei, Bowen Zhang, Lingqiao Liu, Yifan Liu:
ZegCLIP: Towards Adapting CLIP for Zero-shot Semantic Segmentation. 11175-11185
- Luting Wang, Yi Liu , Penghui Du, Zihan Ding, Yue Liao, Qiaosong Qi, Biaolong Chen, Si Liu:
Object-Aware Distillation Pyramid for Open-Vocabulary Object Detection. 11186-11196
- Qingsheng Wang , Lingqiao Liu, Chenchen Jing, Hao Chen, Guoqiang Liang, Peng Wang, Chunhua Shen:
Learning Conditional Attributes for Compositional Zero-Shot Learning. 11197-11206
- Wenbin He, Suphanut Jamonnak, Liang Gou, Liu Ren:
CLIP-S4: Language-Guided Self-Supervised Semantic Segmentation. 11207-11216
- Yanqing Shen, Sanping Zhou, Jingwen Fu, Ruotong Wang, Shitao Chen , Nanning Zheng:
StructVPR: Distill Structural Knowledge with Weighting Samples for Visual Place Recognition. 11217-11226
- Jingyi Zhang, Jiaxing Huang, Xiaoqin Zhang, Shijian Lu:
UniDAformer: Unified Domain Adaptive Panoptic Segmentation Transformer via Hierarchical Mask Calibration. 11227-11237
- Shuting He, Henghui Ding , Wei Jiang:
Primitive Generation and Semantic-Related Alignment for Universal Zero-Shot Segmentation. 11238-11247
- Yuxiang Wei, Zhilong Ji, Xiaohe Wu, Jinfeng Bai, Lei Zhang, Wangmeng Zuo:
Inferring and Leveraging Parts from Object Shape for Improving Semantic Image Synthesis. 11248-11258
- Ju He, Jieneng Chen, Ming-Xian Lin, Qihang Yu, Alan L. Yuille:
Compositor: Bottom-Up Clustering and Compositing for Robust Part and Object Segmentation. 11259-11268
- Sina Hajimiri, Malik Boudiaf, Ismail Ben Ayed, Jose Dolz:
A Strong Baseline for Generalized Few-Shot Semantic Segmentation. 11269-11278
- Ruihuang Li, Chenhang He, Shuai Li, Yabin Zhang, Lei Zhang:
DynaMask: Dynamic Mask Selection for Instance Segmentation. 11279-11288
- Hao Ren, Shoudong Han, Huilin Ding, Ziwen Zhang, Hongwei Wang, Faquan Wang:
Focus On Details: Online Multi-Object Tracking with Diverse Fine-Grained Representation. 11289-11298
- Haoyu He, Jianfei Cai, Zizheng Pan, Jing Liu , Jing Zhang, Dacheng Tao, Bohan Zhuang:
Dynamic Focus-aware Positional Queries for Semantic Segmentation. 11299-11308
- Rohit Jena, Lukas Zhorniyak, Nehal Doiphode, Pratik Chaudhari, Vivek Buch , James C. Gee, Jianbo Shi:
Beyond mAP: Towards Better Evaluation of Instance Segmentation. 11309-11318
- Sun'ao Liu, Yiheng Zhang, Zhaofan Qiu, Hongtao Xie, Yongdong Zhang, Ting Yao:
Learning Orthogonal Prototypes for Generalized Few-Shot Semantic Segmentation. 11319-11328
- Hyekjun Kweon, Sung-Hoon Yoon, Kuk-Jin Yoon:
Weakly Supervised Semantic Segmentation via Adversarial Learning of Classifier and Reconstructor. 11329-11339
- Huimin Huang, Shiao Xie, Lanfen Lin, Ruofeng Tong, Yen-Wei Chen, Yuexiang Li, Hong Wang, Yawen Huang, Yefeng Zheng:
SemiCVT: Semi-Supervised Convolutional Vision Transformer for Semantic Segmentation. 11340-11349
- Zhen Zhao , Lihe Yang, Sifan Long, Jimin Pi, Luping Zhou, Jingdong Wang:
Augmentation Matters: A Simple-Yet-Effective Approach to Semi-Supervised Semantic Segmentation. 11350-11359
- Beomyoung Kim, Joonhyun Jeong, Dongyo Han, Sung Ju Hwang:
The Devil is in the Points: Weakly Semi-Supervised Instance Segmentation via Point-Guided Mask Representation. 11360-11370
- Zilin Luo, Yaoyao Liu , Bernt Schiele, Qianru Sun:
Class-Incremental Exemplar Compression for Class-Incremental Learning. 11371-11380
- Javier Gamazo Tejero, Martin S. Zinkernagel, Sebastian Wolf, Raphael Sznitman, Pablo Márquez-Neila:
Full or Weak Annotations? An Adaptive Strategy for Budget-Constrained Annotation Campaigns. 11381-11391
- Yangyang Shu, Anton van den Hengel, Lingqiao Liu:
Learning Common Rationale to Improve Self-Supervised Representation for Fine-Grained Visual Recognition Problems. 11392-11401
- Lingchen Meng, Xiyang Dai, Yinpeng Chen, Pengchuan Zhang, Dongdong Chen, Mengchen Liu, Jianfeng Wang, Zuxuan Wu, Lu Yuan, Yu-Gang Jiang:
Detection Hub: Unifying Object Detection Datasets via Query Adaptation on Language Embedding. 11402-11411
- Hsin-Ping Huang, Charles Herrmann, Junhwa Hur, Erika Lu, Kyle Sargent, Austin Stone, Ming-Hsuan Yang, Deqing Sun:
Self-supervised AutoFlow. 11412-11421
- Zongheng Tang, Yifan Sun, Si Liu, Yi Yang:
DETR with Additional Global Aggregation for Cross-domain Weakly Supervised Object Detection. 11422-11432
- Zhenyu Wang, Yali Li, Xi Chen, Ser-Nam Lim, Antonio Torralba, Hengshuang Zhao, Shengjin Wang:
Detecting Everything in the Open World: Towards Universal Object Detection. 11433-11443
- Orr Zohar, Kuan-Chieh Wang, Serena Yeung :
PROB: Probabilistic Objectness for Open World Object Detection. 11444-11453



- Yuqing Ma, Hainan Li, Zhange Zhang, Jinyang Guo, Shanghang Zhang, Ruihao Gong, Xianglong Liu: **Annealing-based Label-Transfer Learning for Open World Object Detection.** 11454-11463
- Zihao Wang, Chunxu Wu, Yifei Yang, Zhen Li: **Learning Transformation-Predictive Representations for Detection and Description of Local Features.** 11464-11473
- Muhammad Akhtar Munir, Muhammad Haris Khan, Salman H. Khan, Fahad Shahbaz Khan: **Bridging Precision and Confidence: A Train-Time Loss for Calibrating Object Detection.** 11474-11483
- Mikhail Kennerley, Jian-Gang Wang, Bharadwaj Veeravalli, Robby T. Tan: **2PCNet: Two-Phase Consistency Training for Day-to-Night Unsupervised Domain Adaptive Object Detection.** 11484-11493
- Jiayi Guo, Chaofei Wang, You Wu, Eric J. Zhang, Kai Wang, Xingqian Xu, Shiji Song, Humphrey Shi, Gao Huang: **Zero-Shot Generative Model Adaptation via Image-Specific Prompt Learning.** 11494-11503
- Giacomo Zara, Subhankar Roy, Paolo Rota, Elisa Ricci: **AutoLabel: CLIP-based framework for Open-Set Video Domain Adaptation.** 11504-11513
- Yunhao Bai, Duowen Chen, Qingli Li, Wei Shen, Yan Wang: **Bidirectional Copy-Paste for Semi-Supervised Medical Image Segmentation.** 11514-11524
- Ziyun Yang, Sina Farsiu: **Directional Connectivity-based Segmentation of Medical Images.** 11525-11535
- Aimon Rahman, Jeya Maria Jose Valanarasu, Ilker Hacihaliloglu, Vishal M. Patel: **Ambiguous Medical Image Segmentation Using Diffusion Models.** 11536-11546
- Ramin Nakhli, Puria Azadi Moghadam, Haoyang Mi, Hossein Farahani, Alexander Baras, C. Blake Gilks, Ali Bashashati: **Sparse Multi-Modal Graph Transformer with Shared-Context Processing for Representation Learning of Giga-pixel Images.** 11547-11557
- Zhanyu Wang, Lingqiao Liu, Lei Wang, Luping Zhou: **METransformer: Radiology Report Generation by Transformer with Multiple Learnable Expert Tokens.** 11558-11567
- Siyuan Yan, Zhen Yu, Xuelin Zhang, Dwarikanath Mahapatra, Shekhar S. Chandra, Monika Janda, H. Peter Soyer, Zongyuan Ge: **Towards Trustable Skin Cancer Diagnosis via Rewriting Model's Decision.** 11568-11577
- Jingyao Li, Pengguang Chen, Zexin He, Shaozuo Yu, Shu Liu, Jiaya Jia: **Rethinking Out-of-distribution (OOD) Detection: Masked Image Modeling is All You Need.** 11578-11589
- Ren Wang, Haoliang Sun, Yuling Ma, Xiaoming Xi, Yilong Yin: **MetaViewer: Towards A Unified Multi-View Representation.** 11590-11599
- Jiaqi Jin, Siwei Wang, Zhibin Dong, Xinwang Liu, En Zhu: **Deep Incomplete Multi-View Clustering with Cross-View Partial Sample and Prototype Alignment.** 11600-11609
- Yanglin Feng, Hongyuan Zhu, Dezhong Peng, Xi Peng, Peng Hu: **RONO: Robust Discriminative Learning with Noisy Labels for 2D-3D Cross-Modal Retrieval.** 11610-11619
- Junchi Yu, Jian Liang, Ran He: **Mind the Label Shift of Augmentation-based Graph OOD Generalization.** 11620-11630
- Jinqi Luo, Zhaoning Wang, Chen Henry Wu, Dong Huang, Fernando De la Torre: **Zero-Shot Model Diagnosis.** 11631-11640
- Islam Nassar, Munawar Hayat, Ehsan Abbasnejad, Hamid Rezatofighi, Gholamreza Haffari: **Protocon: Pseudo-Label Refinement via Online Clustering and Prototypical Consistency for Efficient Semi-Supervised Learning.** 11641-11650
- Qi Wei, Lei Feng, Haoliang Sun, Ren Wang, Chenhui Guo, Yilong Yin: **Fine-Grained Classification with Noisy Labels.** 11651-11660
- Zhizhong Huang, Junping Zhang, Hongming Shan: **Twin Contrastive Learning with Noisy Labels.** 11661-11670
- Abhipsa Basu, Sravanti Addepalli, R. Venkatesh Babu: **RMLVQA: A Margin Loss Approach For Visual Question Answering with Language Biases.** 11671-11680
- Jae-Won Cho, Dong-Jin Kim, Hyeonggon Ryu, In So Kweon: **Generative Bias for Robust Visual Question Answering.** 11681-11690
- Ruoyi Du, Dongliang Chang, Kongming Liang, Timothy M. Hospedales, Yi-Zhe Song, Zhanyu Ma: **On-the-Fly Category Discovery.** 11691-11700
- Hugo Touvron, Matthieu Cord, Maxime Oquab, Piotr Bojanowski, Jakob Verbeek, Hervé Jégou: **Co-training 2L Submodels for Visual Recognition.** 11701-11710
- Ruili Feng, Kecheng Zheng, Kai Zhu, Yujun Shen, Jian Zhao, Yukun Huang, Deli Zhao, Jingren Zhou, Michael I. Jordan, Zheng-Jun Zha: **Neural Dependencies Emerging from Learning Massive Categories.** 11711-11720
- Lukas Hoyer, Dengxin Dai, Haoran Wang, Luc Van Gool: **MIC: Masked Image Consistency for Context-Enhanced Domain Adaptation.** 11721-11732



- Dong Zhao, Shuang Wang, Qi Zang, Dou Quan, Xiutiao Ye, Licheng Jiao:
Towards Better Stability and Adaptability: Improve Online Self-Training for Model Adaptation in Semantic Segmentation. 11733-11743
- Ismail Nejjar, Qin Wang, Olga Fink:
DARE-GRAM : Unsupervised Domain Adaptation Regression by Aligning Inverse Gram Matrices. 11744-11754
- Yang Shen, Xuhao Sun, Xiu-Shen Wei:
Equiangular Basis Vectors. 11755-11765
- Mengxi Chen, Linyu Xing, Yu Wang, Ya Zhang:
Enhanced Multimodal Representation Learning with Cross-modal KD. 11766-11775
- Sangrok Lee, Jongseong Bae, Ha Young Kim:
Decompose, Adjust, Compose: Effective Normalization by Playing with Frequency for Domain Generalization. 11776-11785
- Jin Gao , Jialing Zhang, Xihui Liu, Trevor Darrell, Evan Shelhamer, Dequan Wang:
Back to the Source: Diffusion-Driven Adaptation to Test-Time Corruption. 11786-11796
- Shiqi Lin, Zhizheng Zhang, Zhipeng Huang, Yan Lu, Cuiling Lan, Peng Chu, Quanzeng You, Jiang Wang, Zicheng Liu, Amey Parulkar, Viraj Navkal, Zhibo Chen:
Deep Frequency Filtering for Domain Generalization. 11797-11807
- Chiheon Kim, Doyup Lee, Saehoon Kim, Minsu Cho, Wook-Shin Han:
Generalizable Implicit Neural Representations via Instance Pattern Composers. 11808-11817
- Hong-You Chen, Yandong Li, Yin Cui, Mingda Zhang, Wei-Lun Chao, Li Zhang:
Train-Once-for-All Personalization. 11818-11827
- Zitian Chen, Yikang Shen, Mingyu Ding, Zhenfang Chen, Hengshuang Zhao, Erik G. Learned-Miller, Chuang Gan:
Mod-Squad: Designing Mixtures of Experts As Modular Multi-Task Learners. 11828-11837
- Linglan Zhao, Jing Lu, Yunlu Xu, Zhanzhan Cheng, Dashan Guo, Yi Niu, Xiangzhong Fang:
Few-Shot Class-Incremental Learning via Class-Aware Bilateral Distillation. 11838-11847
- Kaiyou Song, Jin Xie, Shan Zhang, Zimeng Luo:
Multi-Mode Online Knowledge Distillation for Self-Supervised Visual Representation Learning. 11848-11857
- Zhiyuan Hu, Yunsheng Li, Jiancheng Lyu, Dashan Gao, Nuno Vasconcelos :
Dense Network Expansion for Class Incremental Learning. 11858-11867
- Ziyou Guo, Haonan Yan, Hui Li, Xiaodong Lin:
Class Attention Transfer Based Knowledge Distillation. 11868-11877
- Yiduo Guo, Bing Liu, Dongyan Zhao:
Dealing with Cross-Task Class Discrimination in Online Continual Learning. 11878-11887
- Yasir Ghunaim , Adel Bibi , Kumail Alhamoud, Motasem Alfarras, Hasan Abed Al Kader Hammoud , Ameya Prabhu, Philip H. S. Torr, Bernard Ghanem :
Real-Time Evaluation in Online Continual Learning: A New Hope. 11888-11897
- Peijie Dong, Lujun Li, Zimian Wei:
DisWOT: Student Architecture Search for Distillation WithOut Training. 11898-11908
- James Seale Smith, Leonid Karlinsky, Vyshnavi Gutta, Paola Cascante-Bonilla, Donghyun Kim, Assaf Arbelle, Rameswar Panda, Rogério Feris, Zsolt Kira:
CODA-Prompt: COntinual Decomposed Attention-Based Prompting for Rehearsal-Free Continual Learning. 11909-11919
- Junha Song, Jungsoo Lee, In So Kweon, Sungha Choi:
EcoTTA: Memory-Efficient Continual Test-Time Adaptation via Self-Distilled Regularization. 11920-11929
- Sanghwan Kim, Lorenzo Noci, Antonio Orvieto, Thomas Hofmann:
Achieving a Better Stability-Plasticity Trade-off via Auxiliary Networks in Continual Learning. 11930-11939
- Shun Lu, Yu Hu, Longxing Yang, Zihao Sun, Jilin Mei, Jianchao Tan, Chengru Song:
PA&DA: Jointly Sampling PAth and DAta for Consistent NAS. 11940-11949
- Lei Zhang, Jie Zhang, Bowen Lei, Subhabrata Mukherjee, Xiang Pan, Bo Zhao, Caiwen Ding, Yao Li, Dongkuan Xu:
Accelerating Dataset Distillation via Model Augmentation. 11950-11959
- Zhaozhi Wang, Kefan Su, Jian Zhang, Huizhu Jia, Qixiang Ye, Xiaodong Xie, Zongqing Lu:
Multi-Agent Automated Machine Learning. 11960-11969
- Erik Gärtner, Luke Metz, Mykhaylo Andriluka, C. Daniel Freeman, Cristian Sminchisescu:
Transformer-Based Learned Optimization. 11970-11979
- Vladimir Kolmogorov:
Solving Relaxations of MAP-MRF Problems: Combinatorial in-Face Frank-Wolfe Directions. 11980-11989
- Jiechao Yang, Yong Liu, Hongteng Xu:
HOTNAS: Hierarchical Optimal Transport for Neural Architecture Search. 11990-12000
- Haechan Noh, Sangeek Hyun, Woojin Jeong, Hanshin Lim, Jae-Pil Heo:
Disentangled Representation Learning for Unsupervised Neural Quantization. 12001-12010



- Guillaume Leclerc, Andrew Ilyas, Logan Engstrom, Sung Min Park, Hadi Salman, Aleksander Madry:
FFCV: Accelerating Training by Removing Data Bottlenecks. 12011-12020
- Jierun Chen, Shiu-Hong Kao , Hao He, Weipeng Zhuo, Song Wen, Chul-Ho Lee, S.-H. Gary Chan:
Run, Don't Walk: Chasing Higher FLOPS for Faster Neural Networks. 12021-12031
- Junho Kim, Byung-Kwan Lee, Yong Man Ro:
Demystifying Causal Features on Adversarial Examples and Causal Inoculation for Robust Network by Adversarial Instrumental Variable Regression. 12032-12042
- Polina Karpikova, Ekaterina Radionova, Anastasia Yaschenko, Andrei Spiridonov, Leonid Kostyushko, Riccardo Fabbricatore, Aleksei Ivakhnenko:
FIANCEE: Faster Inference of Adversarial Networks via Conditional Early Exits. 12032-12043
- Hanjing Wang, Dhiraj Joshi, Shiqiang Wang, Qiang Ji:
Gradient-based Uncertainty Attribution for Explainable Bayesian Deep Learning. 12044-12053
- Xiaotian Yu, Yang Jiang, Tianqi Shi, Zunlei Feng, Yuexuan Wang, Mingli Song, Li Sun:
How to Prevent the Continuous Damage of Noises to Model Training? 12054-12063
- Yongkweon Jeon, Chungman Lee, Ho-Young Kim:
Genie: Show Me the Data for Quantization. 12064-12073
- Fei Zhu, Zhen Cheng, Xu-Yao Zhang, Cheng-Lin Liu:
OpenMix: Exploring Outlier Samples for Misclassification Detection. 12074-12083
- Abhra Chaudhuri , Ayan Kumar Bhunia, Yi-Zhe Song, Anjan Dutta:
Data-Free Sketch-Based Image Retrieval. 12084-12093
- Qingyan Bai, Ceyuan Yang, Yinghao Xu, Xihui Liu, Yujiu Yang , Yujun Shen:
GLeaD: Improving GANs with A Generator-Leading Task. 12094-12104
- Chuangchuang Tan, Yao Zhao, Shikui Wei, Guanghua Gu, Yunchao Wei:
Learning on Gradients: Generalized Artifacts Representation for GAN-Generated Images Detection. 12105-12114
- Hoyoung Choi, Seungwan Jin, Kyungsik Han :
Adversarial Normalization: I Can visualize Everything (ICE). 12115-12124
- Zimeng Zhao, Binghui Zuo, Zhiyu Long, Yangang Wang:
Semi-Supervised Hand Appearance Recovery via Structure Disentanglement and Dual Adversarial Discrimination. 12125-12136
- MyeongAh Cho , Minjung Kim, Sangwon Hwang, Chaewon Park, Kyungjae Lee, Sangyoun Lee:
Look Around for Anomalies: Weakly-Supervised Anomaly Detection via Context-Motion Relational Learning. 12137-12146
- Wenrui Liu, Hong Chang, Bingpeng Ma, Shiguang Shan, Xilin Chen:
Diversity-Measurable Anomaly Detection. 12147-12156
- Yulu Gan, Mingjie Pan, Rongyu Zhang, Zijian Ling, Lingran Zhao, Jiaming Liu, Shanghang Zhang:
Cloud-Device Collaborative Adaptation to Continual Changing Environments in the Real-World. 12157-12166
- Zhe Qu, Xingyu Li, Xiao Han, Rui Duan, Chengchao Shen, Lixing Chen:
How to Prevent the Poor Performance Clients for Personalized Federated Learning? 12167-12176
- Renjie Pi, Weizhong Zhang, Yueqi Xie , Jiahui Gao, Xiaoyu Wang, Sunghun Kim, Qifeng Chen:
DYNAFED: Tackling Client Data Heterogeneity with Global Dynamics. 12177-12186
- Dengsheng Chen, Jie Hu, Vince Junkai Tan, Xiaoming Wei, Enhua Wu:
Elastic Aggregation for Federated Optimization. 12187-12197
- Hideaki Takahashi, Jingjing Liu, Yang Liu:
Breaching FedMD: Image Recovery via Paired-Logits Inversion Attack. 12198-12207
- Tianxin Huang, Zhonggan Ding, Jiangning Zhang, Ying Tai, Zhenyu Zhang, Mingang Chen, Chengjie Wang, Yong Liu:
Learning to Measure the Point Cloud Reconstruction Loss in a Representation Space. 12208-12217
- Lu Pang, Tao Sun, Haibin Ling, Chao Chen:
Backdoor Cleansing with Unlabeled Data. 12218-12227
- Zaixi Zhang, Qi Liu, Zhicai Wang, Zepu Lu, Qingyong Hu:
Backdoor Defense via Deconfounded Representation Learning. 12228-12238
- Ajinkya Tejankar, Maziar Sanjabi, Qifan Wang, Sinong Wang, Hamed Firooz, Hamed Pirsiavash, Liang Tan:
Defending Against Patch-based Backdoor Attacks on Self-Supervised Learning. 12239-12249
- Yi Yu, Yufei Wang, Wenhan Yang, Shijian Lu, Yap-Peng Tan, Alex C. Kot:
Backdoor Attacks Against Deep Image Compression via Adaptive Frequency Trigger. 12250-12259
- Daizong Ding, Erling Jiang, Yuanmin Huang, Mi Zhang, Wenxuan Li, Min Yang:
CAP: Robust Point Cloud Classification via Semantic and Structural Modeling. 12260-12270
- Yang Hou, Qing Guo, Yihao Huang, Xiaofei Xie , Lei Ma, Jianjun Zhao:
Evading DeepFake Detectors via Adversarial Statistical Consistency. 12271-12280
- Zhipeng Wei, Jingjing Chen , Zuxuan Wu, Yu-Gang Jiang:
Enhancing the Self-Universality for Transferable Targeted Attacks. 12281-12290



- Phoenix Neale Williams, Ke Li:
Black-Box Sparse Adversarial Attack via Multi-Objective Optimisation *CVPR Proceedings*. 12291-12301
- Francesco Croce, Sylvestre-Alvise Rebiffé, Evan Shelhamer, Sven Gowal:
Seasoning Model Soups for Robustness to Adversarial and Natural Distribution Shifts. 12313-12323
- Simin Li, Shuning Zhang, Gujun Chen, Dong Wang, Pu Feng, Jiakai Wang, Aishan Liu, Xin Yi, Xianglong Liu:
Towards Benchmarking and Assessing Visual Naturalness of Physical World Adversarial Attacks. 12324-12333
- Xingxing Wei, Jie Yu, Yao Huang:
Physically Adversarial Infrared Patches with Learnable Shapes and Locations. 12334-12342
- Vishal Asnani, Xi Yin, Tal Hassner, Xiaoming Liu:
MaLP: Manipulation Localization Using a Proactive Scheme. 12343-12352
- Daniel S. Jeon, Andreas Meuleman, Seung-Hwan Baek, Min H. Kim:
Polarimetric iToF: Measuring High-Fidelity Depth Through Scattering Media. 12353-12362
- Kun Zhou , Wenbo Li, Yi Wang, Tao Hu, Nianjuan Jiang, Xiaoguang Han, Jiangbo Lu:
NeRFLiX: High-Quality Neural View Synthesis by Learning a Degradation-Driven Inter-viewpoint MiXer. 12363-12374
- Haithem Turki, Jason Y. Zhang, Francesco Ferroni, Deva Ramanan:
SUDS: Scalable Urban Dynamic Scenes. 12375-12385
- Dogyoon Lee, Minhyeok Lee, Chajin Shin, Sangyoun Lee:
DP-NeRF: Deblurred Neural Radiance Field with Physical Scene Priors. 12386-12396
- Heng Yu, Joel Julin, Zoltan A. Milacski, Koichiro Niinuma, László A. Jeni:
DyLiN: Making Light Field Networks Dynamic. 12397-12406
- Ze-Xin Yin, Jiaxiong Qiu, Ming-Ming Cheng, Bo Ren:
Multi-Space Neural Radiance Fields. 12407-12416
- Fernando Rivas-Manzaneque, Jorge Sierra Acosta, Adrián Peñate Sánchez, Francesc Moreno-Noguer, Angela Ribeiro:
NeRFLight: Fast and Light Neural Radiance Fields using a Shared Feature Grid. 12417-12427
- Youngho Yoon, Kuk-Jin Yoon:
Cross-Guided Optimization of Radiance Fields with Multi-View Image Super-Resolution for High-Resolution Novel View Synthesis. 12428-12438
- Jun-Kun Chen, Jipeng Lyu, Yu-Xiong Wang:
NeuralEditor: Editing Neural Radiance Fields via Manipulating Point Clouds. 12439-12448
- Malte Prinzler, Otmar Hilliges, Justus Thies:
DINER: Depth-aware Image-based NEural Radiance fields. 12449-12459
- Agus Gunawan, Soo Ye Kim, Hyeonjun Sim, Jae-Ho Lee , Munchurl Kim:
Modernizing Old Photos Using Multiple References via Photorealistic Style Transfer. 12460-12469
- Xiaoyu Zhang , Yun-Hui Liu:
Efficient Map Sparsification Based on 2D and 3D Discretized Grids. 12470-12478
- Sara Fridovich-Keil, Giacomo Meanti, Frederik Rahbæk Warburg, Benjamin Recht, Angjoo Kanazawa:
K-Planes: Explicit Radiance Fields in Space, Time, and Appearance. 12479-12488
- Jingsen Zhu, Yuchi Huo, Qi Ye, Fujun Luan, Jifan Li, Dianbing Xi, Lisha Wang, Rui Tang, Wei Hua, Hujun Bao, Rui Wang:
I²-SDF: Intrinsic Indoor Scene Reconstruction and Editing via Raytracing in Neural SDFs. 12489-12498
- Zhen Li, Lingli Wang, Mofang Cheng, Cihui Pan, Jiaqi Yang:
Multi-view Inverse Rendering for Large-scale Real-world Indoor Scenes. 12499-12509
- Chenhao Li, Trung Thanh Ngo, Hajime Nagahara:
Inverse Rendering of Translucent Objects using Physical and Neural Renderers. 12510-12520
- Hong-Xing Yu, Samir Agarwala, Charles Herrmann, Richard Szeliski, Noah Snavely, Jiajun Wu, Deqing Sun:
Accidental Light Probes. 12521-12530
- Ruoshi Liu, Carl Vondrick:
Humans as Light Bulbs: 3D Human Reconstruction from Thermal Reflection. 12531-12542
- Suyi Jiang, Haoran Jiang , Ziyu Wang, Haimin Luo, Wenzheng Chen, Lan Xu:
HumanGen: Generating Human Radiance Fields with Explicit Priors. 12543-12554
- Jinguang Tong, Sundaram Muthu , Fahira Afzal Maken, Chuong Nguyen , Hongdong Li:
Seeing Through the Glass: Neural 3D Reconstruction of Object Inside a Transparent Container. 12555-12564
- Thomas P. Ilett, Omer Yuval, Thomas Ranner, Netta Cohen, David C. Hogg:
3D Shape Reconstruction of Semi-Transparent Worms. 12565-12575
- Likang Wang, Lei Chen:
Dionysus: Recovering Scene Structures by Dividing into Semantic Pieces. 12576-12587
- Zhizhuo Zhou, Shubham Tulsiani:
SparseFusion: Distilling View-Conditioned Diffusion for 3D Reconstruction. 12588-12597
- Yiqun Wang, Ivan Skorokhodov, Peter Wonka:
PET-NeuS: Positional Encoding Tri-Planes for Neural Surfaces. 12598-12607



Titas Aciukevicius, Zexiang Xu, Matthew Fisher, Paul Henderson, Hakan Bilen , Niloy J. Mitra, Paul Guerrero:
RenderDiffusion: Image Diffusion for 3D Reconstruction, Inpainting and Generation. 12608-12618

Haochen Wang, Xiaodan Du, Jiahao Li, Raymond A. Yeh, Greg Shakhnarovich:
Score Jacobian Chaining: Lifting Pretrained 2D Diffusion Models for 3D Generation. 12619-12629

Alexander Raistrick, Lahav Lipson, Zeyu Ma, Lingjie Mei, Mingzhe Wang, Yiming Zuo, Karhan Kayan, Hongyu Wen, Beining Han, Yihan Wang, Alejandro Newell, Hei Law, Ankit Goyal, Kaiyu Yang, Jia Deng:
Infinite Photorealistic Worlds Using Procedural Generation. 12630-12641

Muheng Li, Yueqi Duan, Jie Zhou, Jiwen Lu:
Diffusion-SDF: Text-to-Shape via Voxelized Diffusion. 12642-12651

Senmao Li, Joost van de Weijer , Yaxing Wang, Fahad Shahbaz Khan, Meiqin Liu, Jian Yang:
3D-Aware Multi-Class Image-to-Image Translation with NeRFs. 12652-12662

Gal Metzer, Elad Richardson, Or Patashnik, Raja Giryes, Daniel Cohen-Or:
Latent-NeRF for Shape-Guided Generation of 3D Shapes and Textures. 12663-12673

Changwoon Choi , Sang Min Kim, Young Min Kim:
Balanced Spherical Grid for Egocentric View Synthesis. 12663-12673

Junha Hyung, Sungwon Hwang, Daejin Kim, Hyunji Lee, Jaegul Choo:
Local 3D Editing via 3D Distillation of CLIP Knowledge. 12674-12684

Panos Achlioptas, Ian Huang, Minhyuk Sung , Sergey Tulyakov, Leonidas J. Guibas:
ShapeTalk: A Language Dataset and Framework for 3D Shape Edits and Deformations. 12685-12694

Ambareesh Revanur, Debraj Basu, Shradha Agrawal, Dhwanit Agarwal, Deepak Pai:
CoralStyleCLIP: Co-optimized Region and Layer Selection for Image Editing. 12695-12704

Yixuan Li, Chao Ma, Yichao Yan, Wenhan Zhu, Xiaokang Yang:
3D-Aware Face Swapping. 12705-12714

Minchul Kim , Feng Liu, Anil K. Jain, Xiaoming Liu:
DCFace: Synthetic Face Generation with Dual Condition Diffusion Model. 12715-12725

Yujian Zheng, Zirong Jin, Moran Li, Haibin Huang, Chongyang Ma, Shuguang Cui , Xiaoguang Han:
HairStep: Transfer Synthetic to Real Using Strand and Depth Maps for Single-View 3D Hair Modeling. 12726-12735

Zheng Ding, Xuaner Zhang, Zhihao Xia, Lars Jebe, Zhuowen Tu, Xiuming Zhang:
DiffusionRig: Learning Personalized Priors for Facial Appearance Editing. 12736-12746

Libing Zeng, Lele Chen, Wentao Bao, Zhong Li, Yi Xu, Junsong Yuan , Nima K. Kalantari:
3D-aware Facial Landmark Detection via Multi-view Consistent Training on Synthetic Data. 12747-12758

Ricong Huang, Peiwen Lai, Yipeng Qin, Guanbin Li:
Parametric Implicit Face Representation for Audio-Driven Facial Reenactment. 12759-12768

Junxuan Li, Shunsuke Saito, Tomas Simon, Stephen Lombardi, Hongdong Li, Jason M. Saragih:
MEGANE: Morphable Eyeglass and Avatar Network. 12769-12779

Jinbo Xing, Menghan Xia, Yuechen Zhang, Xiaodong Cun, Jue Wang, Tien-Tsin Wong:
CodeTalker: Speech-Driven 3D Facial Animation with Discrete Motion Prior. 12780-12790

Maria-Paola Forte, Peter Kulits, Chun-Hao Huang, Vasileios Choutas, Dimitrios Tzionas , Katherine J. Kuchenbecker, Michael J. Black:
Reconstructing Signing Avatars from Video Using Linguistic Priors. 12791-12801

Korrawe Karunratanakul, Sergey Prokudin, Otmar Hilliges, Siyu Tang :
HARP: Personalized Hand Reconstruction from a Monocular RGB Video. 12802-12813

Hongyi Xu, Guoxian Song, Zihang Jiang, Jianfeng Zhang, Yichun Shi, Jing Liu, Wan-Chun Ma, Jiashi Feng, Linjie Luo:
OmniAvatar: Geometry-Guided Controllable 3D Head Synthesis. 12814-12824

Zhongjin Luo, Shengcai Cai, Jinguo Dong, Ruibo Ming, Liangdong Qiu, Xiaohang Zhan, Xiaoguang Han:
RaBit: Parametric Modeling of 3D Biped Cartoon Characters with a Topological-Consistent Dataset. 12825-12835

Shubh Maheshwari, Rahul Narain , Ramya Hebbalaguppe:
Transfer4D: A Framework for Frugal Motion Capture and Deformation Transfer. 12836-12846

Xingxing Zou, Xintong Han, Waikeung Wong:
CLOTH4D: A Dataset for Clothed Human Reconstruction. 12847-12857

Chen Guo, Tianjian Jiang, Xu Chen, Jie Song, Otmar Hilliges:
Vid2Avatar: 3D Avatar Reconstruction from Videos in the Wild via Self-supervised Scene Decomposition. 12858-12868

Sang-Hun Han, Min-Gyu Park, Ju Hong Yoon, Ju-Mi Kang, Young-Jae Park, Hae-Gon Jeon :
High-fidelity 3D Human Digitization from Single 2K Resolution Images. 12869-12879

Jeonghwan Kim , Mi-Gyeong Gwon, Hyunwoo Park, Hyukmin Kwon, Gi-Mun Um, Wonjun Kim:
Sampling is Matter: Point-Guided 3D Human Mesh Reconstruction. 12880-12889

Zerui Chen, Shizhe Chen, Cordelia Schmid, Ivan Laptev:
gSDF: Geometry-Driven Signed Distance Functions for 3D Hand-Object Reconstruction. 12890-12900



- Boyao Zhou, Di Meng, Jean-Sébastien Franco, Edmond Boyer:
Human Body Shape Completion with Implicit Shape and Flow Learning. 12901-12911
- Zixuan Huang, Varun Jampani, Anh Thai, Yuanzhen Li, Stefan Stojanov, James M. Rehg:
ShapeClipper: Scalable 3D Shape Learning from Single-View Images via Geometric and CLIP-Based Consistency. 12912-12922
- Luke Melas-Kyriazi, Christian Rupprecht, Andrea Vedaldi:
PC²: Projection-Conditioned Point Cloud Diffusion for Single-Image 3D Reconstruction. 12923-12932
- Jiefeng Li, Siyuan Bian, Qi Liu, Jiasheng Tang, Fan Wang, Cewu Lu:
NIKI: Neural Inverse Kinematics with Invertible Neural Networks for 3D Human Pose and Shape Estimation. 12933-12942
- Zicong Fan, Omid Taheri, Dimitrios Tzionas , Muhammed Kocabas, Manuel Kaufmann , Michael J. Black, Otmar Hilliges:
ARCTIC: A Dataset for Dexterous Bimanual Hand-Object Manipulation. 12943-12954
- Zhengdi Yu, Shaoli Huang, Chen Fang, Toby P. Breckon, Jue Wang:
ACR: Attention Collaboration-based Regressor for Arbitrary Two-Hand Reconstruction. 12955-12964
- Hongwei Yi, Chun-Hao P. Huang, Shashank Tripathi, Lea Hering, Justus Thies, Michael J. Black:
MIME: Human-Aware 3D Scene Generation. 12965-12976
- Ming Yan , Xin Wang, Yudi Dai, Siqi Shen, Chenglu Wen, Lan Xu, Yuexin Ma, Cheng Wang:
CIMI4D: A Large Multimodal Climbing Motion Dataset under Human-scene Interactions. 12977-12988
- Zhifeng Lin, Changxing Ding, Huan Yao, Zengsheng Kuang, Shaoli Huang:
Harmonious Feature Learning for Interactive Hand-Object Pose Estimation. 12989-12998
- Takehiko Ohkawa, Kun He, Fadime Sener, Tomas Hodan, Luan Tran, Cem Keskin:
AssemblyHands: Towards Egocentric Activity Understanding via 3D Hand Pose Estimation. 12999-13008
- Haoxuan Qu, Yujun Cai, Lin Geng Foo, Ajay Kumar, Jun Liu:
A Characteristic Function-Based Method for Bottom-Up Human Pose Estimation. 13009-13018
- Lin Geng Foo, Tianjiao Li, Hossein Rahmani, QiuHong Ke, Jun Liu:
Unified Pose Sequence Modeling. 13019-13030
- Jian Wang, Diogo C. Luvizon, Weipeng Xu, Lingjie Liu, Kripasindhu Sarkar, Christian Theobalt:
Scene-Aware Egocentric 3D Human Pose Estimation. 13031-13040
- Jia Gong, Lin Geng Foo, Zhipeng Fan , QiuHong Ke, Hossein Rahmani, Jun Liu:
DiffPose: Toward More Reliable 3D Pose Estimation. 13041-13051
- Jun Chen, Ming Hu, Darren J. Coker, Michael L. Berumen , Blair R. Costelloe, Sara Beery, Anna Rohrbach, Mohamed Elhoseiny :
MammalNet: A Large-Scale Video Benchmark for Mammal Recognition and Behavior Understanding. 13052-13061
- Zifan Shi, Yujun Shen, Yinghao Xu, Sida Peng, Yiyi Liao, Sheng Guo, Qifeng Chen, Dit-Yan Yeung:
Learning 3D-Aware Image Synthesis with Unknown Pose Distribution. 13062-13071
- Yifan Sun, Qixing Huang:
Pose Synchronization under Multiple Pair-wise Relative Poses. 13072-13081
- Can Gümeli, Angela Dai, Matthias Nießner:
ObjectMatch: Robust Registration using Canonical Object Correspondences. 13082-13091
- Anastasis Stathopoulos, Georgios Pavlakos, Ligong Han, Dimitris N. Metaxas:
Learning Articulated Shape with Keypoint Pseudo-Labels from Web Images. 13092-13101
- Dominik Muhle, Lukas Koestler, Krishna Murthy Jatavallabhula, Daniel Cremers:
Learning Correspondence Uncertainty via Differentiable Nonlinear Least Squares. 13102-13112
- Lipu Zhou:
Efficient Second-Order Plane Adjustment. 13113-13121
- Eric Dexheimer, Andrew J. Davison:
Learning a Depth Covariance Function. 13122-13131
- Kunal Chelani, Torsten Sattler, Fredrik Kahl, Zuzana Kukelova:
Privacy-Preserving Representations are not Enough: Recovering Scene Content from Camera Poses. 13132-13141
- Matt Deitke, Dustin Schwenk, Jordi Salvador, Luca Weihs, Oscar Michel, Eli VanderBilt, Ludwig Schmidt, Kiana Ehsani, Aniruddha Kembhavi, Ali Farhadi:
Objaverse: A Universe of Annotated 3D Objects. 13142-13153
- Garrick Brazil, Abhinav Kumar, Julian Straub, Nikhila Ravi, Justin Johnson, Georgia Gkioxari:
Omni3D: A Large Benchmark and Model for 3D Object Detection in the Wild. 13154-13164
- Zhihao Liang, Zhangjin Huang, Changxing Ding, Kui Jia:
HelixSurf: A Robust and Efficient Neural Implicit Surface Learning of Indoor Scenes with Iterative Intertwined Regularization. 13165-13174
- Vojtech Panek, Zuzana Kukelova, Torsten Sattler:
Visual Localization using Imperfect 3D Models from the Internet. 13175-13186



- Yiqing Zhang, Xinming Huang, Ziming Zhang:
PRISE: Demystifying Deep Lucas-Kanade with Strongly Star-Convex Constraints for Multimodel Image Alignment. 13187-13197
- Satoshi Ikehata:
Scalable, Detailed and Mask-Free Universal Photometric Stereo. 13198-13207
- Nishant Jain, Suryansh Kumar , Luc Van Gool:
Enhanced Stable View Synthesis. 13208-13217
- Limeng Qiao, Wenjie Ding, Xi Qiu, Chi Zhang:
End-to-End Vectorized HD-map Construction with Piecewise Bézier Curve. 13218-13228
- Nikita Karaev, Ignacio Rocco, Benjamin Graham, Natalia Neverova, Andrea Vedaldi, Christian Rupprecht:
DynamicStereo: Consistent Dynamic Depth from Stereo Videos. 13229-13239
- Ilya Chugunov, Yuxuan Zhang, Felix Heide:
Shakes on a Plane: Unsupervised Depth Estimation from Unstabilized Photography. 13240-13251
- Stefanie Walz, Mario Bijelic, Andrea Ramazzina, Amanpreet Walia, Fahim Mannan, Felix Heide:
Gated Stereo: Joint Depth Estimation from Gated and Wide-Baseline Active Stereo Cues. 13252-13262
- Yan Yang, Liyuan Pan, Liu Liu, Miaomiao Liu:
K3DN: Disparity-Aware Kernel Estimation for Dual-Pixel Defocus Deblurring. 13263-13272
- Hao Ai, Zidong Cao, Yan-Pei Cao, Ying Shan, Lin Wang :
HRDFuse: Monocular 360° Depth Estimation by Collaboratively Learning Holistic-with-Regional Depth Distributions. 13273-13282
- Fanghua Yu, Xintao Wang, Mingdeng Cao, Gen Li, Ying Shan, Chao Dong:
OSRT: Omnidirectional Image Super-Resolution with Distortion-aware Transformer. 13283-13292
- Hengyi Wang, Jingwen Wang, Lourdes Agapito:
Co-SLAM: Joint Coordinate and Sparse Parametric Encodings for Neural Real-Time SLAM. 13293-13302
- Xintong Liu, Jianyu Wang, Leping Xiao, Xing Fu, Lingyun Qiu, Zuoqiang Shi:
Few-Shot Non-Line-of-Sight Imaging with Signal-Surface Collaborative Regularization. 13303-13312
- Yue Li, Jiayong Peng, Juntian Ye, Yueyi Zhang, Feihu Xu, Zhiwei Xiong:
NLOST: Non-Line-of-Sight Imaging with Transformer. 13313-13322
- Yuto Shibata, Yutaka Kawashima, Mariko Isogawa, Go Irie, Akisato Kimura, Yoshimitsu Aoki:
Listening Human Behavior: 3D Human Pose Estimation with Acoustic Signals. 13323-13332
- Shuo Wang, Xinhai Zhao, Hai-Ming Xu, Zehui Chen, Dameng Yu, Jiahao Chang, Zhen Yang, Feng Zhao:
Towards Domain Generalization for Multi-view 3D Object Detection in Bird-Eye-View. 13333-13342
- Marvin Klingner, Shubhankar Borse, Varun Ravi Kumar, Behnaz Rezaei, Venkatraman Narayanan, Senthil Kumar Yogamani, Fatih Porikli:
X³KD: Knowledge Distillation Across Modalities, Tasks and Stages for Multi-Camera 3D Object Detection. 13343-13353
- Yi Yu, Feipeng Da:
Phase-Shifting Coder: Predicting Accurate Orientation in Oriented Object Detection. 13354-13363
- Anurag Ghosh , N. Dinesh Reddy, Christoph Mertz, Srinivasa G. Narasimhan :
Learned Two-Plane Perspective Prior based Image Resampling for Efficient Object Detection. 13364-13373
- Jinyu Yang, Shang Gao, Zhe Li, Feng Zheng, Ales Leonardis:
Resource-Efficient RGBD Aerial Tracking. 13374-13383
- Ruikang Xu, Chang Chen, Jingyang Peng, Cheng Li, Yibin Huang, Fenglong Song, Youliang Yan, Zhiwei Xiong:
Toward RAW Object Detection: A New Benchmark and A New Model. 13384-13393
- Yingjie Wang, Jiajun Deng, Yao Li, Jinshui Hu, Cong Liu, Yu Zhang, Jianmin Ji, Wanli Ouyang, Yanyong Zhang:
Bi-LRFusion: Bi-Directional LiDAR-Radar Fusion for 3D Dynamic Object Detection. 13394-13403
- Félix Goudreault, Dominik Scheuble, Mario Bijelic, Nicolas Robidoux, Felix Heide:
LiDAR-in-the-Loop Hyperparameter Optimization. 13404-13414
- Martin Büchner, Jannik Zürn, Ion-George Todoran, Abhinav Valada, Wolfram Burgard:
Learning and Aggregating Lane Graphs for Urban Automated Driving. 13415-13424
- Xiaoyan Li, Gang Zhang, Boyue Wang, Yongli Hu, Baocai Yin:
Center Focusing Network for Real-Time LiDAR Panoptic Segmentation. 13425-13434
- Bowei Du, Yecheng Huang, Jiaxin Chen, Di Huang:
Adaptive Sparse Convolutional Networks with Global Context Enhancement for Faster Object Detection on Drone Images. 13435-13444
- Runsen Xu, Tai Wang, Wenwei Zhang, Runjian Chen, Jinkun Cao, Jiangmiao Pang, Dahua Lin:
MV-JAR: Masked Voxel Jigsaw and Reconstruction for LiDAR-Based Self-Supervised Pre-Training. 13445-13454
- Alexandre Boulch, Corentin Sautier, Björn Michele , Gilles Puy, Renaud Marlet:
ALSO: Automotive Lidar Self-Supervision by Occupancy Estimation. 13455-13465



- Shogo Sato, Yasuhiro Yao, Taiga Yoshida, Takuhiro Kaneko, Shingo Ando, Jun Shimamura: **Unsupervised Intrinsic Image Decomposition with LiDAR Intensity.** 13466-13475
- Honghui Yang, Wenxiao Wang, Minghao Chen, Binbin Lin, Tong He, Hua Chen, Xiaofei He, Wanli Ouyang: **PVT-SSD: Single-Stage 3D Object Detector with Point-Voxel Transformer.** 13476-13487
- Yukang Chen, Jianhui Liu, Xiangyu Zhang, Xiaojuan Qi, Jiaya Jia: **LargeKernel3D: Scaling up Kernels in 3D Sparse CNNs.** 13488-13498
- Howard Zhang, Yunhao Ba, Ethan Yang, Varan Mehra, Blake Gella, Akira Suzuki, Arnold Pfahl, Chethan Chinder Chandrappa, Alex Wong , Achuta Kadambi: **WeatherStream: Light Transport Automation of Single Image Deweathering.** 13499-13509
- Ji Hou, Xiaoliang Dai, Zijian He, Angela Dai, Matthias Nießner: **Mask3D: Pretraining 2D Vision Transformers by Learning Masked 3D Priors.** 13510-13519
- Haiyang Wang, Chen Shi, Shaoshuai Shi, Meng Lei, Sen Wang, Di He, Bernt Schiele, Liwei Wang: **DSVT: Dynamic Sparse Voxel Transformer with Rotated Sets.** 13520-13529
- Dasith de Silva Edirimuni , Xuequan Lu , Zhiwen Shao, Gang Li, Antonio Robles-Kelly, Ying He: **IterativePFN: True Iterative Point Cloud Filtering.** 13530-13539
- Hyeon Cho, Junyong Choi, Geonwoo Baek, Wonjun Hwang : **itKD: Interchange Transfer-based Knowledge Distillation for 3D Object Detection.** 13540-13549
- Tuan Duc Ngo, Binh-Son Hua, Khoi Nguyen: **ISBNet: a 3D Point Cloud Instance Segmentation Network with Instance-aware Sampling and Box-aware Dynamic Convolution.** 13550-13559
- Changfeng Ma , Yinuo Chen, Pengxiao Guo, Jie Guo, Chongjun Wang, Yanwen Guo: **Symmetric Shape-Preserving Autoencoder for Unsupervised Real Scene Point Cloud Completion.** 13560-13569
- Xiaoyu Tian, Haoxi Ran , Yue Wang, Hang Zhao: **GeoMAE: Masked Geometric Target Prediction for Self-supervised Point Cloud Pre-Training.** 13570-13580
- Zhikai Chen, Fuchen Long, Zhaofan Qiu, Ting Yao, Wengang Zhou, Jiebo Luo , Tao Mei: **AnchorFormer: Point Cloud Completion from Discriminative Nodes.** 13581-13590
- Qing Li , Huifang Feng, Kanle Shi, Yue Gao, Yi Fang, Yu-Shen Liu, Zhizhong Han: **SHS-Net: Learning Signed Hyper Surfaces for Oriented Normal Estimation of Point Clouds.** 13591-13600
- Xiangyu Zhu, Dong Du , Weikai Chen, Zhiyou Zhao, Yinyu Nie, Xiaoguang Han: **NerVE: Neural Volumetric Edges for Parametric Curve Extraction from Point Cloud.** 13601-13610
- Guofeng Mei , Hao Tang, Xiaoshui Huang, Weijie Wang, Juan Liu, Jian Zhang , Luc Van Gool, Qiang Wu : **Unsupervised Deep Probabilistic Approach for Partial Point Cloud Registration.** 13611-13620
- Junho Shin , Hyo-Jun Lee, Hyunseop Kim, Jong-Hyeon Baek, Daehyun Kim, Yeong Jun Koh: **Local Connectivity-Based Density Estimation for Face Clustering.** 13621-13629
- Tianrui Hui, Zizheng Xun, Fengguang Peng, Junshi Huang, Xiaoming Wei, Xiaolin Wei, Jiao Dai, Jizhong Han, Si Liu: **Bridging Search Region Interaction with Template for RGB-T Tracking.** 13630-13639
- Matteo Farina, Luca Magri, Willi Menapace, Elisa Ricci, Vladislav Golyanik, Federica Arrigoni : **Quantum Multi-Model Fitting.** 13640-13649
- Souhaib Attaiki, Lei Li, Maks Ovsjanikov: **Generalizable Local Feature Pre-training for Deformable Shape Analysis.** 13650-13661
- Jianghao Xiong, Jianhuang Lai: **Similarity Metric Learning For RGB-Infrared Group Re-Identification.** 13662-13671
- Taeyong Song, Sunok Kim, Kwanghoon Sohn: **Unsupervised Deep Asymmetric Stereo Matching with Spatially-Adaptive Self-Similarity.** 13672-13680
- Yikun Bai, Bernhard Schmitzer, Matthew Thorpe, Soheil Kolouri: **Sliced Optimal Partial Transport.** 13681-13690
- Jisoo Jeong, Hong Cai, Risheek Garrepalli, Fatih Porikli: **DistractFlow: Improving Optical Flow Estimation via Realistic Distractions and Pseudo-Labeling.** 13691-13700
- Oleksandr Balabanov, Bernhard Mehlig, Hampus Linander: **Bayesian Posterior Approximation With Stochastic Ensembles.** 13701-13711
- Runsheng Xu, Xin Xia, Jinlong Li, Hanzhao Li , Shuo Zhang, Zhengzhong Tu, Zonglin Meng, Hao Xiang, Xiaoyu Dong, Rui Song , Hongkai Yu, Bolei Zhou, Jiaqi Ma: **V2V4Real: A Real-World Large-Scale Dataset for Vehicle-to-Vehicle Cooperative Perception.** 13712-13722
- Hao Shao, Letian Wang, Ruobing Chen, Steven L. Waslander, Hongsheng Li , Yu Liu: **ReasonNet: End-to-End Driving with Temporal and Global Reasoning.** 13723-13733
- Shaofei Cai, Zihao Wang, Xiaojian Ma, Anji Liu, Yitao Liang: **Open-World Multi-Task Control Through Goal-Aware Representation Learning and Adaptive Horizon Prediction.** 13734-13744
- Luke Rowe, Martin Ethier, Eli-Henry Dykhne, Krzysztof Czarnecki: **FJMP: Factorized Joint Multi-Agent Motion Prediction over Learned Directed Acyclic Interaction Graphs.** 13745-13755



- Davis Rempe, Zhengyi Luo, Xue Bin Peng, Ye Yuan, Kris Kitani, Karsten Kreis, Sanja Fidler, Or Litany: **Trace and Pace: Controllable Pedestrian Animation via Guided Trajectory Diffusion.** 13756-13766
- Vincent-Pierre Berges, Andrew Szot, Devendra Singh Chaplot, Aaron Gokaslan, Roozbeh Mottaghi, Dhruv Batra, Eric Undersander: **Galactic: Scaling End-to-End Reinforcement Learning for Rearrangement at 100k Steps-Per-Second.** 13767-13777
- Guolei Sun, Zhaochong An, Yun Liu, Ce Liu, Christos Sakaridis , Deng-Ping Fan, Luc Van Gool: **Indiscernible Object Counting in Underwater Scenes.** 13791-13801
- Basile Van Hoorick, Pavel Tokmakov, Simon Stent, Jie Li, Carl Vondrick: **Tracking Through Containers and Occluders in the Wild.** 13802-13812
- Jenny Seidenschwarz, Guillem Brasó, Victor Castro Serrano, Ismail Elezi, Laura Leal-Taixé: **Simple Cues Lead to a Strong Multi-Object Tracker.** 13813-13823
- Weijia Li, Saihui Hou, Chunjie Zhang, Chunshui Cao, Xu Liu, Yongzhen Huang, Yao Zhao: **An In-Depth Exploration of Person Re-Identification and Gait Recognition in Cloth-Changing Conditions.** 13824-13833
- Xinqi Fan, Xueli Chen, Mingjie Jiang , Ali Raza Shahid, Hong Yan: **SelfME: Self-Supervised Motion Learning for Micro-Expression Recognition.** 13834-13843
- Jiayu Wang, Kang Zhao, Shiwei Zhang, Yingya Zhang, Yujun Shen, Deli Zhao, Jingren Zhou: **LipFormer: High-fidelity and Generalizable Talking Face Generation with A Pre-learned Facial Codebook.** 13844-13853
- Wenzheng Zeng, Yang Xiao, Sicheng Wei, Jinfang Gan, Xintao Zhang, Zhiguo Cao, Zhiwen Fang, Joey Tianyi Zhou: **Real-time Multi-person Eyeblink Detection in the Wild for Untrimmed Video.** 13854-13863
- Jiaxu Zhang, Junwu Weng, Di Kang, Fang Zhao, Shaoli Huang, Xuefei Zhe, Linchao Bao , Ying Shan, Jue Wang, Zhigang Tu: **Skinned Motion Retargeting with Residual Perception of Motion Semantics & Geometry.** 13864-13872
- Sigal Raab, Inbal Leibovitch, Peizhuo Li, Kfir Aberman, Olga Sorkine-Hornung, Daniel Cohen-Or: **MoDi: Unconditional Motion Synthesis from Diverse Data.** 13873-13883
- Mathias Gehrig, Davide Scaramuzza: **Recurrent Vision Transformers for Object Detection with Event Cameras.** 13884-13893
- Clinton Ansun Mo, Kun Hu, Chengjiang Long, Zhiyong Wang: **Continuous Intermediate Token Learning with Implicit Motion Manifold for Keyframe Based Motion Interpolation.** 13894-13903
- Julius Erbach, Stepan Tulyakov, Patricia Vitoria, Alfredo Bochicchio, Yuanyou Li: **EvShutter: Transforming Events for Unconstrained Rolling Shutter Correction.** 13904-13913
- Jasdeep Singh, Subrahmanyam Murala, G. Sankara Raju Kosuru: **Multi Domain Learning for Motion Magnification.** 13914-13923
- Yixin Yang, Jin Han, Jinxiu Liang, Imari Sato, Boxin Shi: **Learning Event Guided High Dynamic Range Video Reconstruction.** 13924-13934
- Wei Shang, Dongwei Ren, Yi Yang, Hongzhi Zhang, Kede Ma , Wangmeng Zuo: **Joint Video Multi-Frame Interpolation and Deblurring under Unknown Exposure Time.** 13935-13944
- Ce Zheng, Matías Mendieta, Taojiannan Yang, Guo-Jun Qi, Chen Chen: **FeatER: An Efficient Network for Human Reconstruction via Feature Map-Based TransformER.** 13945-13954
- Wenda Zhao, Shigeng Xie, Fan Zhao, You He, Huchuan Lu: **MetaFusion: Infrared and Visible Image Fusion via Meta-Feature Embedding from Object Detection.** 13955-13965
- Shuaizheng Liu, Xindong Zhang, Lingchen Sun, Zhetong Liang, Hui Zeng, Lei Zhang: **Joint HDR Denoising and Fusion: A Real-World Mobile HDR Image Dataset.** 13966-13975
- Muyao Niu, Zhuoxiao Li, Zhihang Zhong, Yinqiang Zheng: **Visibility Constrained Wide-Band Illumination Spectrum Design for Seeing-in-the-Dark.** 13976-13985
- Ji Li, Weixi Wang, Yuesong Nan, Hui Ji: **Self-Supervised Blind Motion Deblurring with Deep Expectation Maximization.** 13986-13996
- Zehua Sheng, Zhu Yu, Xiongwei Liu, Si-Yuan Cao, Yuqi Liu, Hui-Liang Shen, Huaqi Zhang: **Structure Aggregation for Cross-Spectral Stereo Image Guided Denoising.** 13997-14006
- Masakazu Yoshimura, Junji Otsuka, Atsushi Irie, Takeshi Ohashi: **Rawgment: Noise-Accounted RAW Augmentation Enables Recognition in a Wide Variety of Environments.** 14007-14017
- Youssef Mansour, Reinhard Heckel: **Zero-Shot Noise2Noise: Efficient Image Denoising without any Data.** 14018-14027
- Zhaoyang Zhang, Yitong Jiang, Wenqi Shao, Xiaogang Wang, Ping Luo, Kaimo Lin, Jinwei Gu: **Real-Time Controllable Denoising for Image and Video.** 14028-14038
- Zeyu Zhu, Xiangyong Cao, Man Zhou, Junhao Huang, Deyu Meng: **Probability-based Global Cross-modal Upsampling for Pan sharpening.** 14039-14048



- Lanqing Guo, Chong Wang, Wenhan Yang, Siyu Huang , Yufei Wang, Hanspeter Pfister, Bihan Wen
ShadowDiffusion: When Degradation Prior Meets Diffusion Model for Shadow Removal. 14049-14058
- Zizheng Yang, Jie Huang, Jiahao Chang, Man Zhou, Hu Yu, Jinghao Zhang, Feng Zhao:
Visual Recognition-Driven Image Restoration for Multiple Degradation with Intrinsic Semantics Recovery. 14059-14070
- Weixia Zhang, Guangtao Zhai, Ying Wei, Xiaokang Yang, Kede Ma
Blind Image Quality Assessment via Vision-Language Correspondence: A Multitask Learning Perspective. 14071-14081
- Du Chen, Jie Liang, Xindong Zhang, Ming Liu, Hui Zeng, Lei Zhang:
Human Guided Ground-Truth Generation for Realistic Image Super-Resolution. 14082-14091
- Chenyang Qi, Xin Yang , Ka Leong Cheng, Ying-Cong Chen, Qifeng Chen:
Real-time 6K Image Rescaling with Rate-distortion Optimization. 14092-14101
- Jiahao Chao, Zhou Zhou, Hongfan Gao, Jiali Gong, Zhengfeng Yang, Zhenbing Zeng, Lydia Dehbi
Equivalent Transformation and Dual Stream Network Construction for Mobile Image Super-Resolution. 14102-14111
- Yanan Sun, Chi-Keung Tang, Yu-Wing Tai:
Ultrahigh Resolution Image/Video Matting with Spatio-Temporal Sparsity. 14112-14121
- Haiyu Zhao , Yuanbiao Gou, Boyun Li, Dezhong Peng, Jiancheng Lv, Xi Peng:
Comprehensive and Delicate: An Efficient Transformer for Image Restoration. 14122-14132
- Guiwei Zhang, Yongfei Zhang, Tianyu Zhang, Bo Li, Shiliang Pu:
PHA: Patch-Wise High-Frequency Augmentation for Transformer-Based Person Re-Identification. 14133-14142
- Jiarui Lei, Xiaobo Hu, Yue Wang, Dong Liu:
PyramidFlow: High-Resolution Defect Contrastive Localization Using Pyramid Normalizing Flow. 14143-14152
- Zhijie Wu, Yuhe Jin, Kwang Moo Yi:
Neural Fourier Filter Bank. 14153-14163
- Nakkwan Choi, Seungjae Lee, Yongsik Lee, Seungjoon Yang:
Restoration of Hand-Drawn Architectural Drawings using Latent Space Mapping with Degradation Generator. 14164-14172
- Zhanghan Ke, Yuhao Liu , Lei Zhu , Nanxuan Zhao, Rynson W. H. Lau:
Neural Preset for Color Style Transfer. 14173-14182
- Minheng Ni, Xiaoming Li, Wangmeng Zuo:
NÜWA-LIP: Language-guided Image Inpainting with Defect-free VQGAN. 14183-14192
- Ying-Tian Liu , Zhifei Zhang, Yuan-Chen Guo, Matthew Fisher , Zhaowen Wang, Song-Hai Zhang:
DualVector: Unsupervised Vector Font Synthesis with Dual-Part Representation. 14193-14202
- Gwanghyun Kim, Se Young Chun:
DATID-3D: Diversity-Preserved Domain Adaptation Using Text-to-Image Diffusion for 3D Generative Model. 14203-14213
- Ming Tao, Bing-Kun Bao, Hao Tang, Changsheng Xu:
GALIP: Generative Adversarial CLIPs for Text-to-Image Synthesis. 14214-14223
- Dongyeun Lee, Jae Young Lee, Doyeon Kim, Jaehyun Choi, Jaejun Yoo, Junmo Kim:
Fix the Noise: Disentangling Source Feature for Controllable Domain Translation. 14224-14234
- Yuanzhi Zhu, Zhaohai Li, Tianwei Wang, Mengchao He, Cong Yao:
Conditional Text Image Generation with Diffusion Models. 14235-14244
- Zhengyuan Yang, Jianfeng Wang, Zhe Gan, Linjie Li, Kevin Lin, Chenfei Wu, Nan Duan, Zicheng Liu, Ce Liu, Michael Zeng, Lijuan Wang:
ReCo: Region-Controlled Text-to-Image Generation. 14246-14255
- Han Xue, Zhiwu Huang, Qianru Sun, Li Song, Wenjun Zhang:
Freestyle Layout-to-Image Synthesis. 14256-14266
- Haoming Lu, Hazarapet Tunanyan, Kai Wang, Shant Navasardyan, Zhangyang Wang, Humphrey Shi
Specialist Diffusion: Plug-and-Play Sample-Efficient Fine-Tuning of Text-to-Image Diffusion Models to Learn Any Unseen Style. 14267-14276
- Mayu Otani, Riku Togashi, Yu Sawai, Ryosuke Ishigami, Yuta Nakashima, Esa Rahtu , Janne Heikkilä, Shin'ichi Satoh:
Toward Verifiable and Reproducible Human Evaluation for Text-to-Image Generation. 14277-14286
- Naoto Inoue, Kotaro Kikuchi, Edgar Simo-Serra, Mayu Otani, Kota Yamaguchi:
Towards Flexible Multi-modal Document Models. 14287-14296
- Chenlin Meng, Robin Rombach, Ruiqi Gao, Diederik P. Kingma, Stefano Ermon, Jonathan Ho, Tim Salimans:
On Distillation of Guided Diffusion Models. 14297-14306
- Han Zhang, Ruili Feng, Zhantao Yang, Lianghua Huang, Yu Liu, Yifei Zhang, Yujun Shen, Deli Zhao, Jingren Zhou, Fan Cheng:
Dimensionality-Varying Diffusion Process. 14307-14316



- Yao-Chih Lee, Ji-Ze Genevieve Jang, Yi-Ting Chen, Elizabeth Qiu, Jia-Bin Huang:
Shape-Aware Text-Driven Layered Video Editing. 14317-14326
- Yuanbiao Gou, Peng Hu, Jiancheng Lv, Hongyuan Zhu, Xi Peng:
Rethinking Image Super Resolution from Long-Tailed Distribution Learning Perspective. 14327-14336
- Wei-Lun Huang, Ming-Sui Lee:
End-to-end Video Matting with Trimap Propagation. 14337-14347
- Seungmin Jeon, Kwang Pyo Choi, Youngo Park, Chang-Su Kim:
Context-Based Trit-Plane Coding for Progressive Image Compression. 14348-14357
- Zhihao Hu, Dong Xu:
Complexity-guided Slimmable Decoder for Efficient Deep Video Compression. 14358-14367
- Rui Song, Chunyang Fu , Shan Liu, Ge Li:
Efficient Hierarchical Entropy Model for Learned Point Cloud Compression. 14368-14377
- Shishira R. Maiya, Sharath Girish, Max Ehrlich, Hanyu Wang, Kwot Sin Lee, Patrick Poirson, Pengxiang Wu, Chen Wang, Abhinav Shrivastava:
NIRVANA: Neural Implicit Representations of Videos with Adaptive Networks and Autoregressive Patch-Wise Modeling. 14378-14387
- Jinming Liu, Heming Sun, Jiro Katto :
Learned Image Compression with Mixed Transformer-CNN Architectures. 14388-14397
- Jin Lin, Xiaotong Luo, Ming Hong, Yanyun Qu, Yuan Xie, Zongze Wu:
Memory-Friendly Scalable Super-Resolution via Rewinding Lottery Ticket Hypothesis. 14398-14407
- Wenhui Wang, Jifeng Dai, Zhe Chen, Zhenhang Huang, Zhiqi Li, Xizhou Zhu, Xiaowei Hu, Tong Lu, Lewei Lu, Hongsheng Li , Xiaogang Wang, Yu Qiao:
InternImage: Exploring Large-Scale Vision Foundation Models with Deformable Convolutions. 14408-14419
- Xinyu Liu, Houwen Peng, Ningxin Zheng, Yuqing Yang, Han Hu, Yixuan Yuan :
EfficientViT: Memory Efficient Vision Transformer with Cascaded Group Attention. 14420-14430
- Haoran You, Yunyang Xiong, Xiaoliang Dai, Bichen Wu, Peizhao Zhang, Haoqi Fan, Peter Vajda, Yingyan Celine Lin:
Castling-ViT: Compressing Self-Attention via Switching Towards Linear-Angular Attention at Vision Transformer Inference. 14431-14442
- Jiahao Wang, Songyang Zhang, Yong Liu, Taiqiang Wu, Yujiu Yang , Xihui Liu, Kai Chen, Ping Luo, Dahua Lin:
RIFFormer: Keep Your Vision Backbone Effective But Removing Token Mixer. 14443-14452
- Yu Takagi, Shinji Nishimoto:
High-resolution image reconstruction with latent diffusion models from human brain activity. 14453-14463
- Jeremy Speth, Nathan Vance , Patrick J. Flynn, Adam Czajka:
Non-Contrastive Unsupervised Learning of Physiological Signals from Video. 14464-14474
- Zhenda Xie, Zigang Geng, Jingcheng Hu, Zheng Zhang, Han Hu, Yue Cao:
Revealing the Dark Secrets of Masked Image Modeling. 14475-14485
- Samyakh Tukra, Frederick Hoffman, Ken Chatfield:
Improving Visual Representation Learning Through Perceptual Understanding. 14486-14495
- Lucas Beyer, Pavel Izmailov, Alexander Kolesnikov, Mathilde Caron, Simon Kornblith, Xiaohua Zhai, Matthias Minderer, Michael Tschannen, Ibrahim Alabdulmohsin, Filip Pavetic:
FlexiViT: One Model for All Patch Sizes. 14496-14506
- Wele Gedara Chaminda Bandara , Naman Patel, Ali Gholami, Mehdi Nikkhah, Motilal Agrawal, Vishal M. Patel:
AdaMAE: Adaptive Masking for Efficient Spatiotemporal Learning with Masked Autoencoders. 14507-14517
- Chuong Huynh, Yuqian Zhou, Zhe Lin, Connelly Barnes, Eli Shechtman, Sohrab Amirghodsi, Abhinav Shrivastava:
SimpSON: Simplifying Photo Cleanup with Single-Click Distracting Object Segmentation Network. 14518-14527
- Mingyu Ding, Yikang Shen, Lijie Fan, Zhenfang Chen, Zitian Chen, Ping Luo, Joshua B. Tenenbaum, Chuang Gan:
Visual Dependency Transformers: Dependency Tree Emerges from Reversed Attention. 14528-14539
- Alexander Gillert, Giulia Resente, Alba Anadon-Rosell, Martin Wilmking , Uwe Freiherr von Lukas:
Iterative Next Boundary Detection for Instance Segmentation of Tree Rings in Microscopy Images of Shrub Cross Sections. 14540-14548
- Limin Wang, Bingkun Huang, Zhiyu Zhao, Zhan Tong, Yinan He, Yi Wang, Yali Wang, Yu Qiao:
VideoMAE V2: Scaling Video Masked Autoencoders with Dual Masking. 14549-14560
- Qiangqiang Wu , Tianyu Yang, Ziquan Liu, Baoyuan Wu, Ying Shan, Antoni B. Chan:
DropMAE: Masked Autoencoders with Spatial-Attention Dropout for Tracking Tasks. 14561-14571
- Xin Chen, Houwen Peng, Dong Wang, Huchuan Lu, Han Hu:
SeqTrack: Sequence to Sequence Learning for Visual Object Tracking. 14572-14581
- Long Lian, Zhirong Wu, Stella X. Yu:
Bootstrapping Objectness from Videos by Relaxed Common Fate and Visual Grouping. 14582-14591
- Zhiwei Yang, Jing Liu, Zhaoyang Wu, Peng Wu, Xiaotao Liu:
Video Event Restoration Based on Keyframes for Video Anomaly Detection. 14592-14601



- Yucheng Zhao, Chong Luo, Chuanxin Tang, Dongdong Chen, Noel Codella, Zheng-Jun Zha:
Streaming Video Model. 14602-14612
- Jinsheng Xiao, Yuanxu Wu, Yunhua Chen, Shurui Wang, Zhongyuan Wang, Jiayi Ma:
LSTFE-Net: Long Short-Term Feature Enhancement Network for Video Small Object Detection. 14613-14622
- Miran Heo, Sukjun Hwang, Jeongseok Hyun, Hanjung Kim, Seoung Wug Oh, Joon-Young Lee, Seon Joo Kim:
A Generalized Framework for Video Instance Segmentation. 14623-14632
- Dongming Wu, Wencheng Han, Tiancai Wang, Xingping Dong, Xiangyu Zhang, Jianbing Shen:
Referring Multi-Object Tracking. 14633-14642
- Kai Li, Deep Patel, Erik Kruus, Martin Renqiang Min:
Source-Free Video Domain Adaptation with Spatial-Temporal-Historical Consistency Learning. 14643-14652
- Jiadong Wang, Xinyuan Qian, Malu Zhang, Robby T. Tan, Haizhou Li:
Seeing What You Said: Talking Face Generation Guided by a Lip Reading Expert. 14653-14662
- Fiona Ryan, Hao Jiang, Abhinav Shukla, James M. Rehg, Vamsi Krishna Ithapu:
Egocentric Auditory Attention Localization in Conversations. 14663-14674
- Jiaaben Chen, Renrui Zhang, Dongze Lian, Jiaqi Yang, Ziyao Zeng, Jianbo Shi:
iQuery: Instruments as Queries for Audio-Visual Sound Separation. 14675-14686
- Gaoxiang Cong, Liang Li, Yuankai Qi , Zheng-Jun Zha, Qi Wu, Wenyu Wang, Bin Jiang, Ming-Hsuan Yang, Qingming Huang:
Learning to Dub Movies via Hierarchical Prosody Models. 14687-14697
- Madeline Chantry Schiappa, Naman Biyani, Prudvi Kamtam, Shruti Vyas, Hamid Palangi, Vibhav Vineet, Yogesh S. Rawat:
A Large-Scale Robustness Analysis of Video Action Recognition Models. 14698-14708
- Alexandros Stergiou , Dima Damen :
The Wisdom of Crowds: Temporal Progressive Attention for Early Action Prediction. 14709-14719
- Tao Wu , Mengqi Cao, Ziteng Gao, Gangshan Wu, Limin Wang:
STMixer: A One-Stage Sparse Action Detector. 14720-14729
- Jianrong Zhang, Yangsong Zhang, Xiaodong Cun, Yong Zhang, Hongwei Zhao, Hongtao Lu, Xi Shen, Shan Ying:
Generating Human Motion from Textual Descriptions with Discrete Representations. 14730-14740
- Mengyuan Chen, Junyu Gao, Changsheng Xu:
Cascade Evidential Learning for Open-world Weakly-supervised Temporal Action Localization. 14741-14750
- Chen Ju, Kunhao Zheng, Jinxiang Liu, Peisen Zhao, Ya Zhang , Jianlong Chang, Qi Tian, Yanfeng Wang:
Distilling Vision-Language Pre-Training to Collaborate with Weakly-Supervised Temporal Action Localization. 14751-14762
- Jiangwei Lao, Weixiang Hong, Xin Guo , Yingying Zhang, Jian Wang, Jingdong Chen, Wei Chu:
Simultaneously Short- and Long-Term Temporal Modeling for Semi-Supervised Video Semantic Segmentation. 14763-14772
- Difei Gao, Luowei Zhou, Lei Ji, Linchao Zhu, Yi Yang, Mike Zheng Shou:
MIST : Multi-modal Iterative Spatial-Temporal Transformer for Long-form Video Question Answering. 14773-14783
- Daniel McKee, Justin Salamon, Josef Sivic, Bryan C. Russell:
Language-Guided Music Recommendation for Video via Prompt Analogies. 14784-14793
- Yimeng Zhang , Xin Chen, Jinghan Jia, Sijia Liu, Ke Ding:
Text-Visual Prompting for Efficient 2D Temporal Video Grounding. 14794-14804
- Jianhui Yu, Hao Zhu, Liming Jiang, Chen Change Loy, Weidong Cai , Wayne Wu:
CelebV-Text: A Large-Scale Facial Text-Video Dataset. 14805-14814
- Tian Gan, Qing Wang, Xingning Dong, Xiangyuan Ren, Liqiang Nie, Qingpei Guo:
CNVid-3.5M: Build, Filter, and Pre-Train the Large-Scale Public Chinese Video-Text Dataset. 14815-14824
- Yiwu Zhong, Licheng Yu, Yang Bai, Shangwen Li, Xuetong Yan, Yin Li:
Learning Procedure-aware Video Representation from Instructional Videos and Their Narrations. 14825-14835
- Hanlin Wang, Yilu Wu, Sheng Guo, Limin Wang:
PDPP: Projected Diffusion for Procedure Planning in Instructional Videos. 14836-14845
- Xudong Lin, Simran Tiwari, Shiyuan Huang, Manling Li, Mike Zheng Shou, Heng Ji, Shih-Fu Chang:
Towards Fast Adaptation of Pretrained Contrastive Models for Multi-channel Video-Language Retrieval. 14846-14855
- Jingjia Huang, Yinan Li, Jiashi Feng, Xinglong Wu, Xiaoshuai Sun, Rongrong Ji:
Clover: Towards A Unified Video-Language Alignment and Fusion Model. 14856-14866
- Bo He, Jun Wang, Jielin Qiu, Trung Bui, Abhinav Shrivastava, Zhaowen Wang:
Align and Attend: Multimodal Summarization with Dual Contrastive Losses. 14867-14878
- Aisha Urooj Khan, Hilde Kuehne, Bo Wu, Kim Chheu, Walid Bousselham, Chuang Gan, Niels da Vitoria Lobo, Mubarak Shah:
Learning Situation Hyper-Graphs for Video Question Answering. 14879-14889
- Ronglai Zuo, Fangyun Wei, Brian Mak:
Natural Language-Assisted Sign Language Recognition. 14890-14900



- Nikhil Gosala, Kürsat Petek, Paulo L. J. Drews-Jr, Wolfram Burgard, Abhinav Valada: **SkyEye: Self-Supervised Bird's-Eye-View Semantic Mapping Using Monocular Frontal View Images.** 14901-14910
- Chen Gao, Xingyu Peng, Mi Yan, He Wang, Lirong Yang, Haibing Ren, Hongsheng Li , Si Liu: **Adaptive Zone-aware Hierarchical Planner for Vision-Language Navigation.** 14911-14920
- Jacob Krantz, Shurjo Banerjee, Wang Zhu, Jason J. Corso, Peter Anderson, Stefan Lee, Jesse Thomason: **Iterative Vision-and-Language Navigation.** 14921-14930
- Hao Zhu, Raghav Kapoor, So Yeon Min, Winson Han, Jiatai Li, Kaiwen Geng, Graham Neubig, Yonatan Bisk, Aniruddha Kembhavi, Luca Weihs: **EXCALIBUR: Encouraging and Evaluating Embodied Exploration.** 14931-14942
- Yi-Lun Lee, Yi-Hsuan Tsai, Wei-Chen Chiu, Chen-Yu Lee: **Multimodal Prompting with Missing Modalities for Visual Recognition.** 14943-14952
- Tanmay Gupta, Aniruddha Kembhavi: **Visual Programming: Compositional visual reasoning without training.** 14953-14962
- Zhuowan Li, Xingrui Wang, Elias Stengel-Eskin, Adam Kortylewski, Wufei Ma, Benjamin Van Durme, Alan L. Yuille: **Super-CLEVR: A Virtual Benchmark to Diagnose Domain Robustness in Visual Reasoning.** 14963-14973
- Zhenwei Shao, Zhou Yu, Meng Wang, Jun Yu: **Prompting Large Language Models with Answer Heuristics for Knowledge-Based Visual Question Answering.** 14974-14983
- Benjamin Bowman, Alessandro Achille, Luca Zancato, Matthew Trager, Pramuditha Perera, Giovanni Paolini, Stefano Soatto: **À-la-carte Prompt Tuning (APT): Combining Distinct Data Via Composable Prompting.** 14984-14993
- James Seale Smith, Paola Cascante-Bonilla, Assaf Arbelle, Donghyun Kim, Rameswar Panda, David D. Cox, Diyi Yang, Zsolt Kira, Rogério Feris, Leonid Karlinsky: **ConConstruct-VL: Data-Free Continual Structured VL Concepts Learning.** 14994-15004
- Zaid Khan, B. G. Vijay Kumar, Samuel Schulter, Xiang Yu, Yun Fu, Manmohan Chandraker: **Q: How to Specialize Large Vision-Language Models to Data-Scarce VQA Tasks? A: Self-Train on Unlabeled Images!** 15005-15015
- Shruthi Bannur, Stephanie L. Hyland, Qianchu Liu, Fernando Pérez-García, Maximilian Ilse, Daniel C. Castro, Benedikt Boecking, Harshita Sharma , Kenza Bouzid, Anja Thieme, Anton Schwaighofer, Maria Wetscherek, Matthew P. Lungren, Aditya V. Nori, Javier Alvarez-Valle, Ozan Oktay: **Learning to Exploit Temporal Structure for Biomedical Vision-Language Processing.** 15016-15027
- Yunpeng Han, Lisai Zhang , Qingcai Chen, Zhijian Chen, Zhonghua Li, Jianxin Yang, Zhao Cao: **FashionSAP: Symbols and Attributes Prompt for Fine-Grained Fashion Vision-Language Pre-Training.** 15028-15038
- Zhihong Chen, Ruifei Zhang, Yibing Song, Xiang Wan, Guanbin Li: **Advancing Visual Grounding with Scene Knowledge: Benchmark and Method.** 15039-15049
- Weihua Chen, Xianzhe Xu, Jian Jia, Hao Luo, Yaohua Wang, Fan Wang, Rong Jin, Xiuyu Sun: **Beyond Appearance: A Semantic Controllable Self-Supervised Learning Framework for Human-Centric Visual Tasks.** 15050-15061
- Mehdi Zemni, Mickaël Chen, Éloi Zablocki, Hédi Ben-Younes, Patrick Pérez, Matthieu Cord: **OCTET: Object-aware Counterfactual Explanations.** 15062-15071
- Hyesong Choi, Hunsang Lee, Wonil Song, Sangryul Jeon, Kwanghoon Sohn, Dongbo Min: **Local-Guided Global: Paired Similarity Representation for Visual Reinforcement Learning.** 15072-15082
- Pinaki Nath Chowdhury, Ayan Kumar Bhunia, Aneeshan Sain , Subhadeep Koley, Tao Xiang, Yi-Zhe Song : **What Can Human Sketches Do for Object Detection?** 15083-15094
- Yuxiao Chen, Jianbo Yuan, Yu Tian, Shijie Geng, Xinyu Li, Ding Zhou, Dimitris N. Metaxas, Hongxia Yang: **Revisiting Multimodal Representation in Contrastive Learning: From Patch and Token Embeddings to Finite Discrete Tokens.** 15095-15104
- Wei Li , Jiahao Xie, Chen Change Loy: **Correlational Image Modeling for Self-Supervised Visual Pre-Training.** 15105-15115
- Xueyan Zou, Zi-Yi Dou, Jianwei Yang, Zhe Gan, Linjie Li, Chunyuan Li, Xiyang Dai, Harkirat Behl, Jianfeng Wang, Lu Yuan, Nanyun Peng, Lijuan Wang, Yong Jae Lee, Jianfeng Gao: **Generalized Decoding for Pixel, Image, and Language.** 15116-15127
- Cuiqun Chen, Mang Ye , Ding Jiang: **Towards Modality-Agnostic Person Re-identification with Descriptive Query.** 15128-15137
- Hiuyi Cheng, Peirong Zhang, Sihang Wu, Jiaxin Zhang, Qiyuan Zhu, Zecheng Xie, Jing Li, Kai Ding, Lianwen Jin: **M⁶Doc: A Large-Scale Multi-Format, Multi-Type, Multi-Layout, Multi-Language, Multi-Annotation Category Dataset for Modern Document Layout Analysis.** 15138-15147
- Haotian Liu, Kilho Son, Jianwei Yang, Ce Liu, Jianfeng Gao, Yong Jae Lee, Chunyuan Li: **Learning Customized Visual Models with Retrieval-Augmented Knowledge.** 15148-15158
- Zheren Fu, Zhendong Mao, Yan Song, Yongdong Zhang: **Learning Semantic Relationship among Instances for Image-Text Matching.** 15159-15168



- Muhammad Ferjad Naeem, Muhammad Gul Zain Ali Khan, Yongqin Xian, Muhammad Zeshan Afzal, Didier Stricker, Luc Van Gool, Federico Tombari:
I2MVFormer: Large Language Model Generated Multi-View Document Supervision for Zero-Shot Image Classification. 15169-15179
- Rohit Girdhar, Alaaeldin El-Nouby, Zhuang Liu, Mannat Singh, Kalyan Vasudev Alwala, Armand Joulin, Ishan Misra:
ImageBind One Embedding Space to Bind Them All. 15180-15190
- Yusuke Hirota, Yuta Nakashima, Noa Garcia:
Model-Agnostic Gender Debiased Image Captioning. 15191-15200
- Tan Pan, Furong Xu, Xudong Yang, Sifeng He, Chen Jiang , Qingpei Guo, Feng Qian, Xiaobo Zhang, Yuan Cheng, Lei Yang, Wei Chu:
Boundary-aware Backward-Compatible Representation via Adversarial Learning in Image Retrieval. 15201-15210
- Renrui Zhang, Xiangfei Hu, Bohao Li, Siyuan Huang, Hanqiu Deng, Yu Qiao, Peng Gao, Hongsheng Li :
Prompt, Generate, Then Cache: Cascade of Foundation Models Makes Strong Few-Shot Learners. 15211-15222
- Taeho Kil, Seonghyeon Kim, Sukmin Seo, Yoonsik Kim, Daehhee Kim:
Towards Unified Scene Text Spotting Based on Sequence Generation. 15223-15232
- Yanxin Long, Youpeng Wen, Jianhua Han, Hang Xu, Pengzhen Ren , Wei Zhang, Shen Zhao, Xiaodan Liang:
CapDet: Unifying Dense Captioning and Open-World Detection Pretraining. 15233-15243
- Yihan Zeng, Chenhan Jiang, Jiageng Mao, Jianhua Han, Chaoqiang Ye, Qingqiu Huang, Dit-Yan Yeung, Zhen Yang, Xiaodan Liang, Hang Xu:
CLIP²: Contrastive Language-Image-Point Pretraining from Real-World Point Cloud Data. 15244-15253
- Sizhe Wu, Wenwei Zhang, Sheng Jin, Wentao Liu, Chen Change Loy:
Aligning Bag of Regions for Open-Vocabulary Object Detection. 15254-15264
- Chufeng Tang, Lingxi Xie, Xiaopeng Zhang, Xiaolin Hu, Qi Tian:
Visual Recognition by Request. 15265-15274
- Chi Xie , Fangao Zeng, Yue Hu, Shuang Liang, Yichen Wei:
Category Query Learning for Human-Object Interaction Classification. 15275-15284
- Tongkun Guan, Chaochen Gu, Jingzheng Tu, Xue Yang, Qi Feng, Yudi Zhao, Wei Shen:
Self-Supervised Implicit Glyph Attention for Text Recognition. 15285-15294
- Jun Cen, Shiwei Zhang, Xiang Wang, Yixuan Pei, Zhiwu Qing, Yingya Zhang, Qifeng Chen:
Enlarging Instance-specific and Class-specific Information for Open-set Action Recognition. 15295-15304
- Yuqi Lin, Minghao Chen, Wenxiao Wang, Boxi Wu, Ke Li, Binbin Lin, Haifeng Liu, Xiaofei He:
CLIP is Also an Efficient Segmenter: A Text-Driven Approach for Weakly Supervised Semantic Segmentation. 15305-15314
- Shaozhe Hao, Kai Han, Kwan-Yee K. Wong:
Learning Attention as Disentangler for Compositional Zero-Shot Learning. 15315-15324
- Bin Yan, Yi Jiang, Jiannan Wu, Dong Wang, Ping Luo, Zehuan Yuan, Huchuan Lu:
Universal Instance Perception as Object Discovery and Retrieval. 15325-15336
- Man Liu, Feng Li, Chunjie Zhang, Yunchao Wei, Huihui Bai, Yao Zhao:
Progressive Semantic-Visual Mutual Adaption for Generalized Zero-Shot Learning. 15337-15346
- Xiaoxue Chen, Yuhang Zheng, Yupeng Zheng, Qiang Zhou, Hao Zhao, Guyue Zhou, Ya-Qin Zhang:
DPF: Learning Dense Prediction Fields with Weak Supervision. 15347-15357
- Zhibo Yang, Rujiao Long, Pengfei Wang, Sibo Song, Humen Zhong, Wenqing Cheng, Xiang Bai, Cong Yao:
Modeling Entities as Semantic Points for Visual Information Extraction in the Wild. 15358-15367
- Tarun Kalluri, Wangdong Xu, Manmohan Chandraker:
GeoNet: Benchmarking Unsupervised Adaptation across Geographies. 15368-15379
- Maxime Pietrantoni, Martin Humenberger, Torsten Sattler, Gabriela Csurka:
SegLoc: Learning Segmentation-Based Representations for Privacy-Preserving Visual Localization. 15380-15391
- Tai-Yu Pan, Qing Liu, Wei-Lun Chao, Brian L. Price:
Towards Open-World Segmentation of Parts. 15392-15401
- Changdi Yang, Pu Zhao, Yanyu Li, Wei Niu , Jie Xiong Guan, Hao Tang, Minghai Qin, Bin Ren, Xue Lin, Yanzhi Wang:
Pruning Parameterization with Bi-level Optimization for Efficient Semantic Segmentation on the Edge. 15402-15412
- Jian Ding, Nan Xue, Gui-Song Xia , Bernt Schiele, Dengxin Dai:
HGFormer: Hierarchical Grouping Transformer for Domain Generalized Semantic Segmentation. 15413-15423
- Taoseef Ishtiaq, Qing En, Yuhong Guo:
Exemplar-FreeSOLO: Enhancing Unsupervised Instance Segmentation with Exemplars. 15424-15433
- Anurag Das, Yongqin Xian, Dengxin Dai, Bernt Schiele:
Weakly-Supervised Domain Adaptive Semantic Segmentation with Prototypical Contrastive Learning. 15434-15443
- Ying Ji, Yu Wang, Jien Kato:
Spatial-temporal Concept based Explanation of 3D ConvNets. 15444-15453



- Linshan Wu , Zhun Zhong, Leyuan Fang, Xingxin He, Qiang Liu, Jiayi Ma, Hao Chen: **Sparsely Annotated Semantic Segmentation with Adaptive Gaussian Mixtures.** 15454-15464
- Pengchong Qiao, Zhdan Wei, Yu Wang, Zhennan Wang, Guoli Song, Fan Xu, Xiangyang Ji, Chang Liu, Jie Chen: **Fuzzy Positive Learning for Semi-Supervised Semantic Segmentation.** 15465-15474
- Zhenglin Zhou, Huaxia Li, Hong Liu, Nanyang Wang, Gang Yu, Rongrong Ji: **STAR Loss: Reducing Semantic Ambiguity in Facial Landmark Detection.** 15475-15484
- Hao Li , Dingwen Zhang, Nian Liu, Lechao Cheng , Yalun Dai, Chao Zhang, Xinggang Wang, Junwei Han: **Boosting Low-Data Instance Segmentation by Unsupervised Pre-training with Saliency Prompt.** 15485-15494
- Simon Reiß, Constantin Seibold, Alexander Freytag, Erik Rodner, Rainer Stiefelhagen: **Decoupled Semantic Prototypes enable learning from diverse annotation types for semi-weakly segmentation in expert-driven domains.** 15495-15506
- Caixia Zhou, Yaping Huang, Mengyang Pu, Qingji Guan, Li Huang, Haibin Ling: **The Treasure Beneath Multiple Annotations: An Uncertainty-Aware Edge Detector.** 15507-15517
- Tianyu Zhu, Bryce Ferenczi , Pulak Purkait, Tom Drummond, Hamid Rezatofighi, Anton van den Hengel: **Knowledge Combination to Learn Rotated Detection without Rotated Annotation.** 15518-15527
- Xinyi Ying , Li Liu, Yingqian Wang, Ruojing Li , Nuo Chen, Zaiping Lin, Weidong Sheng, Shilin Zhou: **Mapping Degeneration Meets Label Evolution: Learning Infrared Small Target Detection with Single Point Supervision.** 15528-15538
- Yang Liu, Yao Zhang, Yixin Wang, Yang Zhang, Jiang Tian, Zhongchao Shi, Jianping Fan, Zhiqiang He: **SAP-DETR: Bridging the Gap Between Salient Points and Queries-Based Transformer Detector for Fast Model Convergency.** 15539-15547
- Jingyi Xu, Hieu Le, Vu Nguyen, Viresh Ranjan, Dimitris Samaras: **Zero-Shot Object Counting.** 15548-15557
- Wei Hua, Dingkang Liang, Jingyu Li, Xiaolong Liu, Zhikang Zou, Xiaoqing Ye, Xiang Bai: **SOOD: Towards Semi-Supervised Oriented Object Detection.** 15558-15567
- Yue Yao, Tom Gedeon, Liang Zheng: **Large-scale Training Data Search for Object Re-identification.** 15568-15578
- Chang Liu , Weiming Zhang , Xiangru Lin, Wei Zhang, Xiao Tan, Junyu Han, Xiaomao Li, Errui Ding, Jingdong Wang: **Ambiguity-Resistant Semi-Supervised Learning for Dense Object Detection.** 15579-15588
- Shiyu Xia, Jiaqi Lv, Ning Xu, Gang Niu, Xin Geng: **Towards Effective Visual Representations for Partial-Label Learning.** 15589-15598
- Jiakang Yuan, Bo Zhang, Xiangchao Yan, Tao Chen, Botian Shi, Yikang Li, Yu Qiao: **Bi3D: Bi-Domain Active Learning for Cross-Domain 3D Object Detection.** 15599-15608
- Shaokai Wu, Fengyu Yang: **Boosting Detection in Crowd Analysis via Underutilized Output Features.** 15609-15618
- Mahmoud Assran, Quentin Duval, Ishan Misra, Piotr Bojanowski, Pascal Vincent, Michael G. Rabbat, Yann LeCun, Nicolas Ballas: **Self-Supervised Learning from Images with a Joint-Embedding Predictive Architecture.** 15619-15629
- Hongrun Zhang, Liam Burrows, Yanda Meng, Declan Sculthorpe, Abhik Mukherjee, Sarah E. Coupland, Ke Chen , Yalin Zheng: **Weakly Supervised Segmentation with Point Annotations for Histopathology Images via Contrast-Based Variational Model.** 15630-15640
- Hao Jiang, Rushan Zhang, Yanning Zhou, Yumeng Wang, Hao Chen: **DoNet: Deep De-Overlapping Network for Cytology Instance Segmentation.** 15641-15650
- Yongchao Wang, Bin Xiao, Xiuli Bi, Weisheng Li, Xinbo Gao: **MCF: Mutual Correction Framework for Semi-Supervised Medical Image Segmentation.** 15651-15660
- Tsai Hor Chan , Fernando Julio Cendra , Lan Ma, Guosheng Yin, Lequan Yu: **Histopathology Whole Slide Image Analysis with Heterogeneous Graph Representation Learning.** 15661-15670
- Qingjie Zeng, Yutong Xie, Zilin Lu , Yong Xia: **PEFAT: Boosting Semi-Supervised Medical Image Classification via Pseudo-Loss Estimation and Feature Adversarial Training.** 15671-15680
- Xiang Li, Xuelin Qian, Litian Liang, Lingjie Kong, Qiaole Dong , Jiejun Chen, Dingxia Liu, Xiuzhong Yao, Yanwei Fu : **Causally-Aware Intraoperative Imputation for Overall Survival Time Prediction.** 15681-15690
- Hyunjung Choi, Hawook Jeong , Jin Young Choi: **Balanced Energy Regularization Loss for Out-of-distribution Detection.** 15691-15700
- Yeonguk Yu, Sungho Shin, Seongju Lee, Changhyun Jun, Kyoobin Lee: **Block Selection Method for Using Feature Norm in Out-of-Distribution Detection.** 15701-15711
- Jie Wen, Chengliang Liu , Gehui Xu, Zhihao Wu, Chao Huang, Lunke Fei, Yong Xu: **Highly Confident Local Structure Based Consensus Graph Learning for Incomplete Multi-view Clustering.** 15712-15721



- Zeren Chen, Gengshi Huang, Wei Li , Jianing Teng, Kun Wang, Jing Shao, Chen Change Loy, Lu Sheng :
Siamese DETR. 15722-15731
- Xulong Yang, Qing Su, Shihao Ji:
Towards Bridging the Performance Gaps of Joint Energy-Based Models. 15732-15741
- Yun-Hao Cao, Peiqin Sun, Shuchang Zhou:
Three Guidelines You Should Know for Universally Slimmable Self-Supervised Learning. 15742-15751
- Ran Tao, Hao Chen, Marios Savvides:
Boosting Transductive Few-Shot Fine-tuning with Margin-based Uncertainty Weighting and Probability Regularization. 15752-15761
- Jianlong Wu, Haozhe Yang, Tian Gan, Ning Ding, Feijun Jiang, Liqiang Nie:
CHMATCH: Contrastive Hierarchical Matching and Robust Adaptive Threshold Boosted Semi-Supervised Learning. 15762-15772
- Tiberiu Sosea, Cornelia Caragea:
MarginMatch: Improving Semi-Supervised Learning with Pseudo-Margins. 15773-15782
- Kiarash Mohammadi, He Zhao, Mengyao Zhai, Frederick Tung:
Ranking Regularization for Critical Rare Classes: Minimizing False Positives at a High True Positive Rate. 15783-15792
- Zhengzhuo Xu, Ruikang Liu, Shuo Yang, Zenghao Chai, Chun Yuan:
Learning Imbalanced Data with Vision Transformers. 15793-15803
- Yingxiao Du, Jianxin Wu:
No One Left Behind: Improving the Worst Categories in Long-Tailed Learning. 15804-15813
- Fei Du, Peng Yang, Qi Jia, Fengtao Nan, Xiaoting Chen, Yun Yang:
Global and Local Mixture Consistency Cumulative Learning for Long-tailed Visual Recognitions. 15814-15823
- Yanbiao Ma, Licheng Jiao, Fang Liu, Shuyuan Yang, Xu Liu, Lingling Li:
Curvature-Balanced Feature Manifold Learning for Long-Tailed Classification. 15824-15835
- Ping Chen, Xingpeng Zhang , Ye Li, Ju Tao, Bin Xiao, Bing Wang, Zongjie Jiang:
DAA: A Delta Age AdaIN operation for age estimation via binary code transformer. 15836-15845
- Bin Xiao, Yang Hu, Bo Liu, Xiuli Bi, Weisheng Li, Xinbo Gao:
DLBD: A Self-Supervised Direct-Learned Binary Descriptor. 15846-15855
- Tianyun Yang, Danding Wang, Fan Tang, Xinying Zhao, Juan Cao, Sheng Tang:
Progressive Open Space Expansion for Open-Set Model Attribution. 15856-15865
- Fengyi Shen, Akhil Gurram, Ziyuan Liu, He Wang, Alois Knoll:
DiGA: Distil to Generalize and then Adapt for Domain Adaptive Semantic Segmentation. 15866-15877
- Hu Wang, Yuanhong Chen, Congbo Ma, Jodie Avery, Louise Hull, Gustavo Carneiro :
Multi-Modal Learning with Missing Modality via Shared-Specific Feature Modelling. 15878-15887
- Weijie Su, Xizhou Zhu, Chenxin Tao, Lewei Lu, Bin Li, Gao Huang, Yu Qiao, Xiaogang Wang, Jie Zhou, Jifeng Dai:
Towards All-in-One Pre-Training via Maximizing Multi-Modal Mutual Information. 15888-15899
- Xiaorong Qin, Xinhang Song, Shuqiang Jiang:
Bi-Level Meta-Learning for Few-Shot Domain Generalization. 15900-15910
- Luca Zancato, Alessandro Achille, Tian Yu Liu, Matthew Trager, Pramuditha Perera, Stefano Soatto:
Train/Test-Time Adaptation with Retrieval. 15911-15921
- Longhui Yuan, Binhui Xie, Shuang Li:
Robust Test-Time Adaptation in Dynamic Scenarios. 15922-15932
- Yotam Nitzan, Michaël Gharbi, Richard Zhang, Taesung Park, Jun-Yan Zhu, Daniel Cohen-Or, Eli Shechtman:
Domain Expansion of Image Generators. 15933-15942
- Shengsen Wu, Yan Bai, Yihang Lou, Xiongkun Linghu, Jianzhong He, Ling-Yu Duan:
Switchable Representation Learning Framework with Self-Compatibility. 15943-15953
- Hui Tang, Kui Jia:
A New Benchmark: On the Utility of Synthetic Data with Blender for Bare Supervised Learning and Downstream Domain Adaptation. 15954-15964
- Yaoming Wang, Bowen Shi, Xiaopeng Zhang, Jin Li, Yuchen Liu, Wenrui Dai, Chenglin Li, Hongkai Xiong, Qi Tian:
Adapting Shortcut with Normalizing Flow: An Efficient Tuning Framework for Visual Recognition. 15965-15974
- Yulong Tian, Fnu Suya, Anshuman Suri, Fengyuan Xu, David Evans:
Manipulating Transfer Learning for Property Inference. 15975-15984
- Divyam Madaan, Hongxu Yin, Wonmin Byeon, Jan Kautz, Pavlo Molchanov:
Heterogeneous Continual Learning. 15985-15995
- Wei Huang, Zhiliang Peng, Li Dong, Furu Wei, Jianbin Jiao, Qixiang Ye:
Generic-to-Specific Distillation of Masked Autoencoders. 15996-16005
- Yi Xie, Huaidong Zhang, Xuemiao Xu, Jianqing Zhu, Shengfeng He :
Towards a Smaller Student: Capacity Dynamic Distillation for Efficient Image Retrieval. 16006-16015



- Benliu Qiu , Hongliang Li, Haitao Wen, Heqian Qiu, Lanxiao Wang, Fanman Meng, Qingbo Wu, Lili Pan:
CafeBoost: Causal Feature Boost to Eliminate Task-Induced Bias for Class Incremental Learning. 16016-16025
- Xing Nie, Shixiong Xu, Xiyan Liu, Gaofeng Meng, Chunlei Huo, Shiming Xiang:
Bilateral Memory Consolidation for Continual Learning. 16026-16035
- Xingxuan Zhang, Yue He, Renzhe Xu, Han Yu, Zheyen Shen, Peng Cui:
NICO++: Towards Better Benchmarking for Domain Generalization. 16036-16047
- Samyak Jain, Sravanti Addepalli, Pawan Kumar Sahu, Priyam Dey, R. Venkatesh Babu:
DART: Diversify-Aggregate-Repeat Training Improves Generalization of Neural Networks. 16048-16059
- Xuanyang Zhang, Yonggang Li, Xiangyu Zhang, Yongtao Wang, Jian Sun:
Differentiable Architecture Search with Random Features. 16060-16069
- Bingyuan Liu, Jérôme Rony, Adrian Galdran, Jose Dolz, Ismail Ben Ayed:
Class Adaptive Network Calibration. 16070-16079
- Suhyun Kang, Duhun Hwang, Moonjung Eo, Taesup Kim, Wonjong Rhee:
Meta-Learning with a Geometry-Adaptive Preconditioner. 16080-16090
- Gongfan Fang, Xinyin Ma, Mingli Song, Michael Bi Mi, Xinchao Wang:
DepGraph: Towards Any Structural Pruning. 16091-16101
- Zizheng Pan, Jianfei Cai, Bohan Zhuang:
Stitchable Neural Networks. 16102-16112
- Kirill Solodskikh, Azim Kurbanov, Ruslan Aydarkhanov, Irina Zhelavskaya, Yury Parfenov, Dehua Song, Stamatios Lefkimiatis:
Integral Neural Networks. 16113-16122
- Grigoris G. Chrysos, Bohan Wang, Jiankang Deng, Volkan Cevher:
Regularization of polynomial networks for image recognition. 16123-16132
- Sanghyun Woo, Shoubhik Debnath, Ronghang Hu, Xinlei Chen, Zhuang Liu, In So Kweon, Saining Xie:
ConvNeXt V2: Co-designing and Scaling ConvNets with Masked Autoencoders. 16133-16142
- Alexander Binder, Leander Weber, Sebastian Lapuschkin , Grégoire Montavon, Klaus-Robert Müller , Wojciech Samek :
Shortcomings of Top-Down Randomization-Based Sanity Checks for Evaluations of Deep Neural Network Explanations. 16143-16152
- Thomas Fel, Melanie Ducoffe, David Vigouroux, Rémi Cadène, Mikael Capelle, Claire Nicodème, Thomas Serre:
Don't Lie to Me! Robust and Efficient Explainability with Verified Perturbation Analysis. 16153-16163
- Chuanwen Feng, Yilong Ren, Xike Xie:
OT-Filter: An Optimal Transport Filter for Learning with Noisy Labels. 16164-16174
- Zhuo Huang, Miaozi Zhu, Xiaobo Xia, Li Shen, Jun Yu, Chen Gong, Bo Han, Bo Du, Tongliang Liu:
Robust Generalization Against Photon-Limited Corruptions via Worst-Case Sharpness Minimization. 16175-16185
- Yuanpeng Tu, Boshen Zhang, Yuxi Li, Liang Liu, Jian Li, Jiangning Zhang, Yabiao Wang, Chengjie Wang, Cairong Zhao:
Learning with Noisy labels via Self-supervised Adversarial Noisy Masking. 16186-16195
- Chen Lin, Bo Peng, Zheyang Li, Wenming Tan, Ye Ren, Jun Xiao, Shiliang Pu:
Bit-shrinking: Limiting Instantaneous Sharpness for Improving Post-training Quantization. 16196-16205
- Jongheon Jeong , Sihyun Yu, Hankook Lee, Jinwoo Shin:
Enhancing Multiple Reliability Measures via Nuisance-Extended Information Bottleneck. 16206-16218
- Haozhe Liu, Wentian Zhang, Bing Li, Haoqian Wu, Nanjun He, Yawen Huang, Yuexiang Li, Bernard Ghanem , Yefeng Zheng:
AdaptiveMix: Improving GAN Training via Feature Space Shrinkage. 16219-16229
- Divya Saxena, Jiannong Cao, Jiahao Xu, Tarun Kulshrestha :
Re-GAN: Data-Efficient GANs Training via Architectural Reconfiguration. 16230-16240
- Yang Liu, Shen Yan, Laura Leal-Taixé, James Hays, Deva Ramanan:
Soft Augmentation for Image Classification. 16241-16250
- Zhaodi Zhang, Zhiyi Xue, Yang Chen, Si Liu , Yueling Zhang, Jing Liu, Min Zhang :
Boosting Verified Training for Robust Image Classifications via Abstraction. 16251-16260
- Reza Akbarian Bafghi, Danna Gurari:
A New Dataset Based on Images Taken by Blind People for Testing the Robustness of Image Classification Models Trained for ImageNet Categories. 16261-16270
- Chen Zhang , Guorong Li, Yuankai Qi , Shuhui Wang, Laiyun Qing, Qingming Huang, Ming-Hsuan Yang:
Exploiting Completeness and Uncertainty of Pseudo Labels for Weakly Supervised Video Anomaly Detection. 16271-16280
- Hui Zhang, Zuxuan Wu, Zheng Wang , Zheneng Chen, Yu-Gang Jiang:
Prototypical Residual Networks for Anomaly Detection and Localization. 16281-16291
- Ming Li, Qingli Li, Yan Wang:
Class Balanced Adaptive Pseudo Labeling for Federated Semi-Supervised Learning. 16292-16301



- Meirui Jiang, Holger R. Roth, Wenqi Li, Dong Yang, Can Zhao, Vishwesh Nath, Daguang Xu, Qi Dou, Ziyue Xu:
Fair Federated Medical Image Segmentation via Client Contribution Estimation. 16302-16311
- Wenke Huang, Mang Ye, Zekun Shi, He Li , Bo Du:
Rethinking Federated Learning with Domain Shift: A Prototype View. 16312-16322
- Yuanhao Xiong, Ruochen Wang, Minhaoo Cheng , Felix Yu, Cho-Jui Hsieh:
FedDM: Iterative Distribution Matching for Communication-Efficient Federated Learning. 16323-16332
- Hagay Michaeli, Tomer Michaeli, Daniel Soudry:
Alias-Free Convnets: Fractional Shift Invariance via Polynomial Activations. 16333-16342
- Ka Ho Chow, Ling Liu, Wenqi Wei, Fatih Ilhan , Yanzhao Wu:
STDLens: Model Hijacking-Resilient Federated Learning for Object Detection. 16343-16351
- Shiwei Feng, Guanhong Tao, Siyuan Cheng, Guangyu Shen, Xiangzhe Xu, Yingqi Liu, Kaiyuan Zhang, Shiqing Ma, Xiangyu Zhang:
Detecting Backdoors in Pre-trained Encoders. 16352-16362
- Xiaogeng Liu, Minghui Li , Haoyu Wang, Shengshan Hu, Dengpan Ye, Hai Jin, Libing Wu, Chaowei Xiao:
Detecting Backdoors During the Inference Stage Based on Corruption Robustness Consistency. 16363-16372
- Zeyang Sha, Xinlei He, Ning Yu, Michael Backes, Yang Zhang :
Can't Steal? Cont-Steal! Contrastive Stealing Attacks Against Image Encoders. 16373-16383
- Ngoc-Bao Nguyen, Keshigeyan Chandrasegaran, Milad Abdollahzadeh, Ngai-Man Cheung:
Re-Thinking Model Inversion Attacks Against Deep Neural Networks. 16384-16393
- Binghui Wang, Meng Pang, Yun Dong:
Turning Strengths into Weaknesses: A Certified Robustness Inspired Attack Framework against Graph Neural Networks. 16394-16403
- Weiwei Feng, Nanqing Xu, Tianzhu Zhang, Yongdong Zhang:
Dynamic Generative Targeted Attacks with Pattern Injection. 16404-16414
- Jianping Zhang, Yizhan Huang, Weibin Wu, Michael R. Lyu:
Transferable Adversarial Attacks on Vision Transformers with Token Gradient Regularization. 16415-16424
- Guillaume Jeanneret, Loïc Simon, Frédéric Jurie:
Adversarial Counterfactual Visual Explanations. 16425-16435
- Ziquan Liu, Yi Xu, Xiangyang Ji, Antoni B. Chan:
TWINS: A Fine-Tuning Framework for Improved Transferability of Adversarial Robustness and Generalization. 16436-16446
- Gaojie Jin, Xinping Yi, Dengyu Wu, Ronghui Mu, Xiaowei Huang:
Randomized Adversarial Training via Taylor Expansion. 16447-16457
- Zifan Wang, Nan Ding, Tomer Levinboim, Xi Chen, Radu Soricut:
Improving Robust Generalization by Direct PAC-Bayesian Bound Minimization. 16458-16468
- Fahad Shamshad, Koushik Srivatsan, Karthik Nandakumar:
Evading Forensic Classifiers with Attribute-Conditioned Adversarial Faces. 16469-16478
- Baowei Jiang, Bing Bai, Haozhe Lin, Yu Wang, Yuchen Guo, Lu Fang:
DartBlur: Privacy Preservation with Detection Artifact Suppression. 16479-16488
- Tomoki Ichikawa, Yoshiki Fukao, Shohei Nobuhara, Ko Nishino:
Fresnel Microfacet BRDF: Unification of Polari-Radiometric Surface-Body Reflection. 16489-16497
- Xiaomeng Xu, Yanchao Yang, Kaichun Mo, Boxiao Pan, Li Yi, Leonidas J. Guibas:
JacobiNeRF: NeRF Shaping with Mutual Information Gradients. 16498-16507
- Hao Yang, Lanqing Hong, Aoxue Li, Tianyang Hu, Zhenguo Li, Gim Hee Lee, Liwei Wang:
ContraNeRF: Generalizable Neural Radiance Fields for Synthetic-to-real Novel View Synthesis via Contrastive Learning. 16508-16517
- Mikaela Angelina Uy, Ricardo Martin-Brualla, Leonidas J. Guibas, Ke Li:
SCADE: NeRFs from Space Carving with Ambiguity-Aware Depth Estimates. 16518-16527
- Silvan Weder, Guillermo Garcia-Hernando, Áron Monszpart, Marc Pollefeys, Gabriel J. Brostow, Michael Firman, Sara Vicente:
Removing Objects From Neural Radiance Fields. 16528-16538
- Andreas Meuleman, Yu-Lun Liu, Chen Gao, Jia-Bin Huang, Changil Kim, Min H. Kim, Johannes Kopf:
Progressively Optimized Local Radiance Fields for Robust View Synthesis. 16539-16548
- Chen Yang , Peihao Li, Zanwei Zhou, Shanxin Yuan, Bingbing Liu, Xiaokang Yang, Weichao Qiu, Wei Shen:
NeRFVS: Neural Radiance Fields for Free View Synthesis via Geometry Scaffolds. 16549-16558
- Zhe Jun Tang, Tat-Jen Cham, Haiyu Zhao:
ABLE-NeRF: Attention-Based Rendering with Learnable Embeddings for Neural Radiance Field. 16559-16568
- Zhiqin Chen, Thomas A. Funkhouser, Peter Hedman, Andrea Tagliasacchi:
MobileNeRF: Exploiting the Polygon Rasterization Pipeline for Efficient Neural Field Rendering on Mobile Architectures. 16569-16578



- Henry Peters, Yunhao Ba, Achuta Kadambi:
pCON: Polarimetric Coordinate Networks for Neural Scene Representations. 16579-16589
- Siqi Yang, Xuanning Cui, Yongjie Zhu, Jiajun Tang , Si Li, Zhaofei Yu, Boxin Shi:
Complementary Intrinsic Properties from Neural Radiance Fields and CNNs for Outdoor Scene Relighting. 16600-16609
- Benjamin Attal, Jia-Bin Huang, Christian Richardt, Michael Zollhöfer, Johannes Kopf, Matthew O'Toole, Changil Kim:
HyperReel: High-Fidelity 6-DoF Video with Ray-Conditioned Sampling. 16610-16620
- Yue Chen, Xuan Wang, Xingyu Chen, Qi Zhang, Xiaoyu Li, Yu Guo, Jue Wang, Fei Wang:
UV Volumes for Real-time Rendering of Editable Free-view Human Performance. 16621-16631
- Ruizhi Shao, Zerong Zheng, Hanzhang Tu, Boning Liu , Hongwen Zhang , Yebin Liu:
Tensor4D: Efficient Neural 4D Decomposition for High-Fidelity Dynamic Reconstruction and Rendering. 16632-16642
- Yichen Sheng, Jianming Zhang, Julien Philip, Yannick Hold-Geoffroy, Xin Sun, He Zhang, Lu Ling, Bedrich Benes:
PixHt-Lab: Pixel Height Based Light Effect Generation for Image Compositing. 16643-16653
- Sepideh Sarajian Maralan, Chris Careaga, Yagiz Aksoy :
Computational Flash Photography through Intrinsic Properties. 16654-16662
- Shun Iwase, Shunsuke Saito, Tomas Simon, Stephen Lombardi, Timur M. Bagautdinov, Rohan Joshi, Fabian Prada, Takaaki Shiratori, Yaser Sheikh, Jason M. Saragih:
RelightableHands: Efficient Neural Relighting of Articulated Hand Models. 16663-16673
- Jaehoon Choi, Dongki Jung, Taejae Lee, Sangwook Kim, Youngdong Jung, Dinesh Manocha, Donghwan Lee:
TMO: Textured Mesh Acquisition of Objects with a Mobile Device by using Differentiable Rendering. 16674-16684
- Yufan Ren, Fangjinhua Wang, Tong Zhang, Marc Pollefeys, Sabine Süsstrunk:
VolRecon: Volume Rendering of Signed Ray Distance Functions for Generalizable Multi-View Reconstruction. 16685-16695
- Pierre Zins, Yuanlu Xu, Edmond Boyer, Stefanie Wuhrer, Tony Tung:
Multi-View Reconstruction Using Signed Ray Distance Functions (SRDF). 16696-16706
- Mingfang Zhang , Jinglu Wang , Xiao Li, Yifei Huang, Yoichi Sato, Yan Lu:
Structural Multiplane Image: Bridging Neural View Synthesis and 3D Reconstruction. 16707-16716
- Chamin Hewa Koneputugodage, Yizhak Ben-Shabat, Stephen Gould:
Octree Guided Unoriented Surface Reconstruction. 16717-16726
- Xianghui Yang, Guosheng Lin, Zhenghao Chen, Luping Zhou:
Neural Vector Fields: Implicit Representation by Explicit Learning. 16727-16738
- Richard Liu, Noam Aigerman, Vladimir G. Kim, Rana Hanocka:
DA Wand: Distortion-Aware Selection Using Neural Mesh Parameterization. 16739-16749
- Siyuan Huang, Zan Wang, Puhal Li, Baoxiong Jia, Tengyu Liu, Yixin Zhu, Wei Liang, Song-Chun Zhu:
Diffusion-based Generation, Optimization, and Planning in 3D Scenes. 16750-16761
- Weiyu Li, Xuelin Chen, Jue Wang, Baoquan Chen:
Patch-Based 3D Natural Scene Generation from a Single Example. 16762-16772
- Hung-Yu Tseng, Qinbo Li, Changil Kim, Suhib Alsisan, Jia-Bin Huang, Johannes Kopf:
Consistent View Synthesis with Pose-Guided Diffusion Models. 16773-16783
- Yuhan Li, Yishun Dou, Xuanhong Chen, Bingbing Ni, Yilin Sun, Yutian Liu, Fuzhen Wang:
Generalized Deep 3D Shape Prior via Part-Discretized Diffusion Process. 16784-16794
- Tianyu Luan, Yuanhao Zhai, Jingjing Meng, Zhong Li, Zhang Chen, Yi Xu, Junsong Yuan:
High Fidelity 3D Hand Shape Reconstruction via Scalable Graph Frequency Decomposition. 16795-16804
- Jiacheng Wei, Hao Wang , Jiashi Feng, Guosheng Lin, Kim-Hui Yap:
TAPS3D: Text-Guided 3D Textured Shape Generation from Pseudo Supervision. 16805-16815
- Pu Li, Jianwei Guo, Xiaopeng Zhang, Dong-Ming Yan:
SECAD-Net: Self-Supervised CAD Reconstruction by Learning Sketch-Extrude Operations. 16816-16826
- Namhyuk Ahn, Patrick Kwon, Jihye Back, Kibeom Hong, Seungkwon Kim:
Interactive Cartoonization with Controllable Perceptual Factors. 16827-16835
- Dejan Azinovic, Olivier Maury, Christophe Hery, Matthias Nießner, Justus Thies:
High-Res Facial Appearance Capture from Polarized Smartphone Images. 16836-16846
- Richard Plesh, Peter Peer, Vitomir Struc:
GlassesGAN: Eyewear Personalization Using Synthetic Appearance Discovery and Targeted Subspace Modeling. 16847-16857
- Prashanth Chandran, Gaspard Zoss, Paulo F. U. Gotardo, Derek Bradley:
Continuous Landmark Detection with 3D Queries. 16858-16867
- Mingwu Zheng, Haiyu Zhang, Hongyu Yang, Di Huang:
NeuFace: Realistic 3D Neural Face Rendering from Multi-View Images. 16868-16877
- Aggelina Chatziagapi, Dimitris Samaras:
AVFace: Towards Detailed Audio-Visual 4D Face Reconstruction. 16878-16889



- Ziqian Bai, Feitong Tan, Zeng Huang, Kripasindhu Sarkar, Danhang Tang, Di Qiu, Abhimitra Meka, Ruofei Du, Mingsong Dou, Sergio Orts-Escalano, Rohit Pandey, Ping Tan, Thabo Beeler, Sean Fanello, Yinda Zhang:
Learning Personalized High Quality Volumetric Head Avatars from Monocular RGB Videos. 16890-16900
- Zhiyuan Ma, Xiangyu Zhu, Guojun Qi, Zhen Lei, Lei Zhang:
OTAvatar: One-Shot Talking Face Avatar with Controllable Tri-Plane Rendering. 16901-16910
- Kaiyue Shen, Chen Guo, Manuel Kaufmann , Juan Jose Zarate, Julien Valentin, Jie Song, Otmar Hilliges:
X-Avatar: Expressive Human Avatars. 16911-16921
- Tianjian Jiang, Xu Chen, Jie Song, Otmar Hilliges:
InstantAvatar: Learning Avatars from Monocular Video in 60 Seconds. 16922-16932
- Xi Wang, Robin Courant, Jinglei Shi, Éric Marchand, Marc Christie:
JAWS: Just A Wild Shot for Cinematic Transfer in Neural Radiance Fields. 16933-16942
- Zhengming Yu , Wei Cheng, Xian Liu, Wayne Wu, Kwan-Yee Lin:
MonoHuman: Animatable Human Neural Field from Monocular Video. 16943-16953
- Enric Corona, Mihai Zanfir, Thiemo Alldieck, Eduard Gabriel Bazavan, Andrei Zanfir, Cristian Sminchisescu:
Structured 3D Features for Reconstructing Controllable Avatars. 16954-16964
- Artur Grigorev, Michael J. Black, Otmar Hilliges:
HOOD: Hierarchical Graphs for Generalized Modelling of Clothing Dynamics. 16965-16974
- Zhanhao Hu, Wenda Chu, Xiaopei Zhu, Hui Zhang, Bo Zhang, Xiaolin Hu:
Physically Realizable Natural-Looking Clothing Textures Evade Person Detectors via 3D Modeling. 16975-16984
- Xiaokun Sun, Qiao Feng, Xiongzheng Li, Jinsong Zhang, Yu-Kun Lai, Jingyu Yang, Kun Li:
Learning Semantic-Aware Disentangled Representation for Flexible 3D Human Body Editing. 16985-16994
- Gengshan Yang, Chaoyang Wang, N. Dinesh Reddy, Deva Ramanan:
Reconstructing Animatable Categories from Videos. 16995-17005
- Yusuke Yoshiyasu:
Deformable Mesh Transformer for 3D Human Mesh Recovery. 17006-17015
- Yifei Yin, Chen Guo, Manuel Kaufmann , Juan Jose Zarate, Jie Song, Otmar Hilliges:
Hi4D: 4D Instance Segmentation of Close Human Interaction. 17016-17027
- Gyeongsik Moon:
Bringing Inputs to Shared Domains for 3D Interacting Hands Recovery in the Wild. 17028-17037
- Zehong Shen, Zhi Cen, Sida Peng, Qing Shuai, Hujun Bao, Xiaowei Zhou:
Learning Human Mesh Recovery in 3D Scenes. 17038-17047
- Hao Xu, Tianyu Wang, Xiao Tang, Chi-Wing Fu:
H2ONet: Hand-Occlusion-and-Orientation-Aware Network for Real-Time 3D Hand Mesh Reconstruction. 17048-17058
- Ruoshi Liu, Sachit Menon, Chengzhi Mao, Dennis Park, Simon Stent, Carl Vondrick:
What You Can Reconstruct from a Shadow. 17059-17068
- Yu Ren, Ronghan Chen, Yang Cong:
Autonomous Manipulation Learning for Similar Deformable Objects via Only One Demonstration. 17069-17078
- Shreyas Hampali, Tomas Hodan, Luan Tran, Lingni Ma, Cem Keskin, Vincent Lepetit:
In-Hand 3D Object Scanning from an RGB Sequence. 17079-17088
- Sumith Kulal, Tim Brooks, Alex Aiken, Jiajun Wu, Jimei Yang, Jingwan Lu, Alexei A. Efros, Krishna Kumar Singh:
Putting People in Their Place: Affordance-Aware Human Insertion into Scenes. 17089-17099
- Yixin Chen, Sai Kumar Dwivedi, Michael J. Black, Dimitrios Tzionas :
Detecting Human-Object Contact in Images. 17100-17110
- Zitian Tang, Wenjie Ye, Wei-Chiu Ma, Hang Zhao:
What Happened 3 Seconds Ago? Inferring the Past with Thermal Imaging. 17111-17120
- Xiaogang Peng, Siyuan Mao, Zizhao Wu:
Trajectory-Aware Body Interaction Transformer for Multi-Person Pose Forecasting. 17121-17130
- Runyang Feng, Yixing Gao, Xueqing Ma, Tze Ho Elden Tse, Hyung Jin Chang:
Mutual Information-Based Temporal Difference Learning for Human Pose Estimation in Video. 17131-17141
- Jiaman Li, C. Karen Liu , Jiajun Wu:
Ego-Body Pose Estimation via Ego-Head Pose Estimation. 17142-17151
- Jeeseung Park, Jin-Woo Park, Jong-Seok Lee:
ViPLO: Vision Transformer Based Pose-Conditioned Self-Loop Graph for Human-Object Interaction Detection. 17152-17162
- Linfang Zheng , Chen Wang, Yinghan Sun, Esha Dasgupta, Hua Chen, Ales Leonardis, Wei Zhang, Hyung Jin Chang:
HS-Pose: Hybrid Scope Feature Extraction for Category-level Object Pose Estimation. 17163-17173
- Chen Li, Gim Hee Lee:
ScarceNet: Animal Pose Estimation with Scarce Annotations. 17174-17183



- Qixia Lin, Linlin Yang, Angela Yao:
Cross-Domain 3D Hand Pose Estimation with Dual Modalities. 17184-17193
- Keyu Yan, Tingwei Gao, Hui Zhang, Chengjun Xie:
Linking Garment with Person via Semantically Associated Landmarks for Virtual Try-On. 17194-17204
- Yuxi Xiao, Nan Xue, Tianfu Wu , Gui-Song Xia :
Level-S²fM: Structure from Motion on Neural Level Set of Implicit Surfaces. 17205-17214
- Ganlin Zhang , Viktor Larsson, Daniel Barath:
Revisiting Rotation Averaging: Uncertainties and Robust Losses. 17215-17224
- Ted de Vries Lentsch, Zimin Xia , Holger Caesar , Julian F. P. Kooij :
SliceMatch: Geometry-Guided Aggregation for Cross-View Pose Estimation. 17225-17234
- Liyan Chen, Weihan Wang, Philippas Mordohai:
Learning the Distribution of Errors in Stereo Matching for Joint Disparity and Uncertainty Estimation. 17235-17244
- Shen Yan, Yu Liu, Long Wang, Zehong Shen, Zhen Peng, Haomin Liu , Maojun Zhang, Guofeng Zhang, Xiaowei Zhou:
Long-Term Visual Localization with Mobile Sensors. 17245-17255
- Nilesh Kulkarni, Linyi Jin, Justin Johnson, David F. Fouhey:
Learning to Predict Scene-Level Implicit 3D from Posed RGBD Data. 17256-17265
- Chunghwan Lee, Jaihoon Kim, Chanhyuk Yun, Je Hyeong Hong :
Paired-Point Lifting for Enhanced Privacy-Preserving Visual Localization. 17266-17275
- Ruohan Gao, Yiming Dou, Hao Li, Tanmay Agarwal, Jeannette Bohg , Yunzhu Li, Li Fei-Fei, Jiajun Wu:
The Object Folder Benchmark : Multisensory Learning with Neural and Real Objects. 17276-17286
- Tianyu Huang, Haoang Li, Kejing He, Congying Sui, Bin Li, Yun-Hui Liu:
Learning Accurate 3D Shape Based on Stereo Polarimetric Imaging. 17287-17296
- Mehdi S. M. Sajjadi, Aravindh Mahendran, Thomas Kipf, Etienne Pot, Daniel Duckworth, Mario Lucic, Klaus Greff:
RUST: Latent Neural Scene Representations from Unposed Imagery. 17297-17306
- Linyi Jin, Jianming Zhang, Yannick Hold-Geoffroy, Oliver Wang, Kevin Blackburn-Matzen, Matthew Sticha, David F. Fouhey:
Perspective Fields for Single Image Camera Calibration. 17307-17316
- Huiyu Gao, Wei Mao, Miaomiao Liu:
VisFusion: Visibility-Aware Online 3D Scene Reconstruction from Videos. 17317-17326
- Rémi Pautrat, Daniel Barath, Viktor Larsson, Martin R. Oswald , Marc Pollefeys:
DeepLSD: Line Segment Detection and Refinement with Deep Image Gradients. 17327-17336
- Zhijie Shen, Zishuo Zheng, Chunyu Lin, Lang Nie, Kang Liao, Shuai Zheng, Yao Zhao:
Disentangling Orthogonal Planes for Indoor Panoramic Room Layout Estimation with Cross-Scale Distortion Awareness. 17337-17345
- Ce Liu, Suryansh Kumar , Shuhang Gu, Radu Timofte , Luc Van Gool:
Single Image Depth Prediction Made Better: A Multivariate Gaussian Take. 17346-17356
- Qi Zhang, Hongdong Li, Qing Wang:
Wide-Angle Rectification via Content-Aware Conformal Mapping. 17357-17365
- Hanyue Lou, Minggui Teng, Yixin Yang, Boxin Shi:
All-in-Focus Imaging from Event Focal Stack. 17366-17375
- Yisu Zhang, Jianke Zhu, Lixiang Lin:
Multi-View Stereo Representation Revist: Region-Aware MVSNet. 17376-17385
- Fangfu Liu, Chubin Zhang, Yu Zheng, Yueqi Duan:
Semantic Ray: Learning a Generalizable Semantic Field with Cross-Reprojection Attention. 17386-17396
- Weijia Li, Yawen Lai, Linning Xu, Yuanbo Xiangli, Jinhua Yu, Conghui He, Gui-Song Xia , Dahua Lin:
OmniCity: Omnipotent City Understanding with Multi-Level and Multi-View Images. 17397-17407
- Mohammad Mahdi Johari, Camilla Carta, François Fleuret:
ESLAM: Efficient Dense SLAM System Based on Hybrid Representation of Signed Distance Fields. 17408-17419
- Jianyu Wang, Xintong Liu, Leping Xiao, Zuqiang Shi, Lingyun Qiu, Xing Fu:
Non-Line-of-Sight Imaging with Signal Superresolution Network. 17420-17429
- Mohammed Alloulah, Maximilian Arnold:
Look, Radiate, and Learn: Self-Supervised Localisation via Radio-Visual Correspondence. 17430-17440
- Vudit Vudit, Martin Engelberge, Mathieu Salzmann:
Learning Transformations to Reduce the Geometric Shift in Object Detection. 17441-17450
- Shaofei Huang, Zhenwei Shen, Zehao Huang, Zi-han Ding, Jiao Dai, Jizhong Han, Naiyan Wang, Si Liu:
Anchor3DLane: Learning to Regress 3D Anchors for Monocular 3D Lane Detection. 17451-17460
- Xiaowei Chi, Jiaming Liu, Ming Lu, Rongyu Zhang, Zhaoqing Wang, Yandong Guo, Shanghang Zhang:
BEV-SAN: Accurate BEV 3D Object Detection via Slice Attention Networks. 17461-17470
- Wenhao Wu , Hau-San Wong, Si Wu:
Semi-Supervised Stereo-Based 3D Object Detection via Cross-View Consensus. 17471-17481



- Runzhou Tao, Wencheng Han, Zhongying Qiu, Cheng-Zhong Xu, Jianbing Shen:
Weakly Supervised Monocular 3D Object Detection Using Multi-View Projection and Direction Consistency. 17482-17492
- Yunsong Zhou, Hongzi Zhu, Quan Liu, Shan Chang, Minyi Guo:
MonoATT: Online Monocular 3D Object Detection with Adaptive Token Transformer. 17493-17503
- Yu-Jhe Li, Shawn Hunt, Jinyung Park, Matthew O'Toole, Kris Kitani:
Azimuth Super-Resolution for FMCW Radar in Autonomous Driving. 17504-17513
- Xindi Wu, KwunFung Lau, Francesco Ferroni, Aljosa Osep, Deva Ramanan:
Pix2Map: Cross-Modal Retrieval for Inferring Street Maps from Images. 17514-17523
- Xin Li, Tao Ma, Yuenan Hou, Botian Shi, Yuchen Yang, Youquan Liu, Xingjiao Wu , Qin Chen, Yikang Li, Yu Qiao, Liang He:
LoGoNet: Towards Accurate 3D Object Detection with Local-to-Global Cross-Modal Fusion. 17524-17534
- Xuan Xiong, Yicheng Liu, Tianyuan Yuan, Yue Wang, Yilun Wang, Hang Zhao:
Neural Map Prior for Autonomous Driving. 17535-17544
- Xin Lai, Yukang Chen, Fanbin Lu, Jianhui Liu, Jiaya Jia:
Spherical Transformer for LiDAR-Based 3D Recognition. 17545-17555
- Qianjiang Hu , Daizong Liu, Wei Hu:
Density-Insensitive Unsupervised Domain Adaption on 3D Object Detection. 17556-17566
- Jinyu Li, Chenxu Luo, Xiaodong Yang:
PillarNeXt: Rethinking Network Designs for 3D Object Detection in LiDAR Point Clouds. 17567-17576
- Liwen Zhang, Xinyan Zhang, Youcheng Zhang, Yufei Guo, Yuanpei Chen, Xuhui Huang, Zhe Ma:
PeakConv: Learning Peak Receptive Field for Radar Semantic Segmentation. 17577-17586
- Hyeonseong Kim, Yoonsu Kang, Changgyoon Oh, Kuk-Jin Yoon:
Single Domain Generalization for LiDAR Semantic Segmentation. 17587-17598
- Ruibo Li, Hanyu Shi, Ziang Fu, Zhe Wang, Guosheng Lin:
Weakly Supervised Class-agnostic Motion Prediction for Autonomous Driving. 17599-17608
- Satish Kumar, Ivan Arevalo, A. S. M. Iftekhar, B. S. Manjunath:
MethaneMapper: Spectral Absorption Aware Hyperspectral Transformer for Methane Detection. 17609-17618
- Zihui Zhang, Bo Yang, Bing Wang, Bo Li:
GrowSP: Unsupervised Semantic Segmentation of 3D Point Clouds. 17619-17629
- Yushuang Wu, Zizheng Yan, Ce Chen , Lai Wei, Xiao Li, Guanbin Li, Yihao Li, Shuguang Cui , Xiaoguang Han:
SCoDA: Domain Adaptive Shape Completion for Real Scans. 17630-17641
- Zhaoyang Xia, Youquan Liu, Xin Li, Xinge Zhu, Yuexin Ma, Yikang Li, Yuenan Hou, Yu Qiao:
SCPNet: Semantic Scene Completion on Point Cloud. 17642-17651
- Jiajing Chen, Minmin Yang, Senem Velipasalar:
ViewNet: A Novel Projection-Based Backbone with View Pooling for Few-shot Point Cloud Classification. 17652-17660
- Zhuoyang Zhang, Yuhao Dong, Yunze Liu, Li Yi:
Complete-to-Partial 4D Distillation for Self-Supervised Point Cloud Sequence Representation Learning. 17661-17670
- Cheng Wen, Baosheng Yu, Dacheng Tao:
Learnable Skeleton-Aware 3D Point Cloud Sampling. 17671-17681
- Haojia Lin, Xiawu Zheng, Lijiang Li, Fei Chao, Shanshan Wang, Yan Wang, Yonghong Tian, Rongrong Ji:
Meta Architecture for Point Cloud Analysis. 17682-17691
- Hehe Fan, Linchao Zhu, Yi Yang, Mohan S. Kankanhalli:
PointListNet: Deep Learning on 3D Point Lists. 17692-17701
- Junle Yu, Luwei Ren, Wenhui Zhou , Yu Zhang, Lili Lin, Guojun Dai:
PEAL: Prior-embedded Explicit Attention Learning for Low-overlap Point Cloud Registration. 17702-17711
- Chao Chen, Yu-Shen Liu, Zhizhong Han:
Unsupervised Inference of Signed Distance Functions from Single Sparse Point Clouds without Learning Priors. 17712-17723
- Baorui Ma, Junsheng Zhou, Yu-Shen Liu, Zhizhong Han:
Towards Better Gradient Consistency for Neural Signed Distance Functions via Level Set Alignment. 17724-17734
- Dongliang Cao, Florian Bernard:
Self-Supervised Learning for Multimodal Non-Rigid 3D Shape Matching. 17735-17744
- Xiyu Zhang, Jiaqi Yang, Shikun Zhang, Yanning Zhang:
3D Registration with Maximal Cliques. 17745-17754
- Zhixin Ling, Zhen Xing, Xiangdong Zhou, Manliang Cao, Guichun Zhou:
PanoSwin: a Pano-style Swin Transformer for Panorama Understanding. 17755-17764
- Johan Edstedt, Ioannis Athanasiadis, Mårten Wadenbäck, Michael Felsberg:
DKM: Dense Kernelized Feature Matching for Geometry Estimation. 17765-17775



- Junjie Ni, Yijin Li, Zhaoyang Huang, Hongsheng Li , Hujun Bao, Zhaopeng Cui, Guofeng Zhang:
PATS: Patch Area Transportation with Subdivision for Local Feature Matching. 17776-17786
- Yixuan Sun, Dongyang Zhao, Zhangyue Yin, Yiwen Huang, Tao Gui, Wenqiang Zhang, Weifeng Ge:
Correspondence Transformers with Asymmetric Feature Learning and Matching Flow Super-Resolution. 17787-17796
- Hoonhee Cho, Jegyeong Cho, Kuk-Jin Yoon:
Learning Adaptive Dense Event Stereo from the Image Domain. 17797-17807
- Liangzu Peng, Christian Kümmerle, René Vidal:
On the Convergence of IRLS and Its Variants in Outlier-Robust Estimation. 17808-17818
- Jie Hu, Linyan Huang, Tianhe Ren, Shengchuan Zhang, Rongrong Ji, Liujuan Cao:
You Only Segment Once: Towards Real-Time Panoptic Segmentation. 17819-17829
- Chenyu Yang , Yuntao Chen, Hao Tian, Chenxin Tao, Xizhou Zhu, Zhaoxiang Zhang, Gao Huang, Hongyang Li, Yu Qiao, Lewei Lu, Jie Zhou, Jifeng Dai:
BEVFormer v2: Adapting Modern Image Backbones to Bird's-Eye-View Recognition via Perspective Supervision. 17830-17839
- Yuanzheng Ci, Yizhou Wang, Meilin Chen, Shixiang Tang, Lei Bai, Feng Zhu, Rui Zhao, Fengwei Yu, Donglian Qi, Wanli Ouyang:
UniHCP: A Unified Model for Human-Centric Perceptions. 17840-17852
- Yihan Hu, Jiazheng Yang, Li Chen, Keyu Li, Chonghao Sima, Xizhou Zhu, Siqi Chai, Senyao Du, Tianwei Lin, Wenhui Wang, Lewei Lu, Xiaosong Jia, Qiang Liu, Jifeng Dai, Yu Qiao, Hongyang Li:
Planning-oriented Autonomous Driving. 17853-17862
- Zikang Zhou , Jianping Wang, Yung-Hui Li, Yu-Kai Huang:
Query-Centric Trajectory Prediction. 17863-17873
- Guangyi Chen , Zhenhao Chen, Shunxing Fan, Kun Zhang:
Unsupervised Sampling Promoting for Stochastic Human Trajectory Prediction. 17874-17884
- Hyung-Gun Chi , Kwonjoon Lee, Nakul Agarwal, Yi Xu, Karthik Ramani, Chiho Choi:
AdamsFormer for Spatial Action Localization in the Future. 17885-17895
- Ram Ramrakhya, Dhruv Batra, Erik Wijmans, Abhishek Das:
PIRLNav: Pretraining with Imitation and RL Finetuning for OBJECTNAV. 17896-17906
- Allan Zhou, Moo Jin Kim, Lirui Wang, Pete Florence, Chelsea Finn :
NeRF in the Palm of Your Hand: Corrective Augmentation for Robotics via Novel-View Synthesis. 17907-17917
- Naisong Luo, Yuwen Pan, Rui Sun, Tianzhu Zhang, Zhiwei Xiong, Feng Wu:
Camouflaged Instance Segmentation via Explicit De-Camouflaging. 17918-17927
- Ziqi Pang, Jie Li, Pavel Tokmakov, Dian Chen, Sergey Zagoruyko, Yu-Xiong Wang:
Standing Between Past and Future: Spatio-Temporal Modeling for Multi-Camera 3D Multi-Object Tracking. 17928-17938
- Zheng Qin, Sanping Zhou, Le Wang, Jinghai Duan, Gang Hua, Wei Tang:
MotionTrack: Learning Robust Short-Term and Long-Term Motions for Multi-Object Tracking. 17939-17948
- Yufeng Cui, Yimei Kang:
Multi-modal Gait Recognition via Effective Spatial-Temporal Feature Fusion. 17949-17957
- Hanyang Wang, Bo Li, Shuang Wu, Siyuan Shen, Feng Liu, Shouhong Ding, Aimin Zhou:
Rethinking the Learning Paradigm for Dynamic Facial Expression Recognition. 17958-17968
- Weichuang Li, Longhao Zhang, Dong Wang, Bin Zhao, Zhigang Wang, Mulin Chen, Bang Zhang, Zhongjian Wang, Liefeng Bo, Xuelong Li:
One-Shot High-Fidelity Talking-Head Synthesis with Deformable Neural Radiance Field. 17969-17978
- Duomin Wang, Yu Deng, Zixin Yin, Heung-Yeung Shum, Baoyuan Wang:
Progressive Disentangled Representation Learning for Fine-Grained Controllable Talking Head Synthesis. 17979-17989
- Chengzhi Cao, Xueyang Fu, Hongjian Liu, Yukun Huang, Kunyu Wang, Jiebo Luo , Zheng-Jun Zha:
Event-Guided Person Re-Identification via Sparse-Dense Complementary Learning. 17990-17999
- Xin Chen , Biao Jiang, Wen Liu, Zilong Huang, Bin Fu, Tao Chen, Gang Yu:
Executing your Commands via Motion Diffusion in Latent Space. 18000-18010
- Xiang Wang , Shiwei Zhang, Zhiwu Qing, Changxin Gao, Yingya Zhang, Deli Zhao, Nong Sang:
MoLo: Motion-Augmented Long-Short Contrastive Learning for Few-Shot Action Recognition. 18011-18021
- Lexuan Xu, Guang Hua, Haijian Zhang, Lei Yu, Ning Qiao:
"Seeing" Electric Network Frequency from Events. 18022-18031
- Taewoo Kim, Yujeong Chae, Hyun-Karl Jang, Kuk-Jin Yoon:
Event-based Video Frame Interpolation with Cross-Modal Asymmetric Bidirectional Motion Fields. 18032-18042
- Lei Sun, Christos Sakaridis , Jingyun Liang, Peng Sun, Kai Zhang, Jiezhang Cao, Qi Jiang, Kaiwei Wang, Luc Van Gool:
Event-Based Frame Interpolation with Ad-hoc Deblurring. 18043-18052



- Jiaqi Xu, Xiaowei Hu, Lei Zhu, Qi Dou, Jifeng Dai, Yu Qiao, Pheng-Ann Heng: **Video Dehazing via a Multi-Range Temporal Alignment Network with Physical Prior.** 18053-18062
- Yawen Lu, Qifan Wang, Siqi Ma, Tong Geng, Yingjie Victor Chen, Huaijin G. Chen, Dongfang Liu: **TransFlow: Transformer as Flow Learner.** 18063-18073
- Hao Zhang, Feng Li, Huaizhe Xu, Shijia Huang, Shilong Liu, Lionel M. Ni, Lei Zhang: **MP-Former: Mask-Piloted Transformer for Image Segmentation.** 18074-18083
- Lin Tian, Thomas Hastings Greer, François-Xavier Vialard, Roland Kwitt, Raúl San José Estépar, Richard Garrett Rushmore, Nikolaos Makris, Sylvain Bouix, Marc Niethammer: **GradICON: Approximate Diffeomorphisms via Gradient Inverse Consistency.** 18084-18094
- Yang Zhou, Kaijian Chen, Rongjun Xiao, Hui Huang: **Neural Texture Synthesis with Guided Correspondence.** 18095-18104
- Zhenxuan Fang, Fangfang Wu, Weisheng Dong, Xin Li, Jinjian Wu, Guangming Shi: **Self-supervised Non-uniform Kernel Estimation with Flow-based Motion Prior for Blind Image Deblurring.** 18105-18114
- Yang Wang, Long Peng, Liang Li, Yang Cao, Zheng-Jun Zha: **Decoupling-and-Aggregating for Image Exposure Correction.** 18115-18124
- Huiyuan Fu, Wenkai Zheng, Xiangyu Meng, Xin Wang , Chuanming Wang, Huadong Ma: **You Do Not Need Additional Priors or Regularizers in Retinex-Based Low-Light Image Enhancement.** 18125-18134
- Xin Jin, Linghao Han, Zhen Li, Chun-Le Guo, Zhi Chai, Chongyi Li: **DNF: Decouple and Feedback Network for Seeing in the Dark.** 18135-18144
- Shirui Huang, Keyan Wang, Huan Liu, Jun Chen, Yunsong Li: **Contrastive Semi-Supervised Learning for Underwater Image Restoration via Reliable Bank.** 18145-18155
- Zichun Wang, Ying Fu, Ji Liu, Yulun Zhang: **LG-BPN: Local and Global Blind-Patch Network for Self-Supervised Real-World Denoising.** 18156-18165
- Tao Liu, Jun Cheng , Shan Tan: **Spectral Bayesian Uncertainty for Image Super-Resolution.** 18166-18175
- Taihui Li, Hengkang Wang, Zhong Zhuang, Ju Sun: **Deep Random Projector: Accelerated Deep Image Prior.** 18176-18185
- Chao Wang, Zhedong Zheng , Ruijie Quan, Yifan Sun, Yi Yang: **Context-Aware Pretraining for Efficient Blind Image Decomposition.** 18186-18195
- Leyi Li, Huijie Qiao, Qi Ye, Qinmin Yang: **Metadata-Based RAW Reconstruction via Implicit Neural Functions.** 18196-18205
- Yufei Wang, Yi Yu, Wenhan Yang, Lanqing Guo, Lap-Pui Chau, Alex C. Kot, Bihan Wen : **Raw Image Reconstruction with Learned Compact Metadata.** 18206-18215
- Juncheol Ye, Hyunho Yeo, Jinwoo Park, Dongsu Han: **AccelIR: Task-aware Image Compression for Accelerating Neural Restoration.** 18216-18226
- Ziwen Chen, Kaushik Patnaik, Shuangfei Zhai, Alvin Wan, Zhile Ren, Alexander G. Schwing, Alex Colburn, Fuxin Li: **AutoFocusFormer: Image Segmentation off the Grid.** 18227-18236
- Nando Metzger, Rodrigo Caye Daudt, Konrad Schindler: **Guided Depth Super-Resolution by Deep Anisotropic Diffusion.** 18237-18246
- Min Wei, Xuesong Zhang: **Super-Resolution Neural Operator.** 18247-18256
- Hao-Wei Chen, Yu-Syuan Xu, Min-Fong Hong, Yi-Min Tsai, Hsien-Kai Kuo, Chun-Yi Lee: **Cascaded Local Implicit Transformer for Arbitrary-Scale Super-Resolution.** 18257-18267
- Hoang M. Le, Brian L. Price, Scott Cohen, Michael S. Brown: **GamutMLP: A Lightweight MLP for Color Loss Recovery.** 18268-18277
- Yawei Li , Yuchen Fan, Xiaoyu Xiang, Denis Demandolx, Rakesh Ranjan, Radu Timofte , Luc Van Gool: **Efficient and Explicit Modelling of Image Hierarchies for Image Restoration.** 18278-18289
- Sheng Liu, Cong Phuoc Huynh, Cong Chen, Maxim Arap, Raffay Hamid: **LEMaRT: Label-Efficient Masked Region Transform for Image Harmonization.** 18290-18299
- Linfeng Wen, Chengying Gao, Changqing Zou: **CAP-VSTNet: Content Affinity Preserved Versatile Style Transfer.** 18300-18309
- Yizhi Song, Zhifei Zhang, Zhe L. Lin, Scott Cohen, Brian L. Price, Jianming Zhang, Soo Ye Kim, Daniel G. Aliaga: **ObjectStitch: Object Compositing with Diffusion Model.** 18310-18319
- Yuqing Wang, Yizhi Wang, Longhui Yu, Yuesheng Zhu, Zhouhui Lian: **DeepVecFont-v2: Exploiting Transformers to Synthesize Vector Fonts with Higher Quality.** 18320-18328
- Hao Tang , Songhua Liu, Tianwei Lin, Shaoli Huang, Fu Li, Dongliang He, Xinchao Wang: **Master: Meta Style Transformer for Controllable Zero-Shot and Few-Shot Artistic Style Transfer.** 18329-18338
- Aditya Sanghi, Rao Fu, Vivian Liu, Karl D. D. Willis, Hooman Shayani, Amir Hosein Khasahmadi, Srinath Sridhar, Daniel Ritchie:

**CLIP-Sculptor: Zero-Shot Generation of High-Fidelity and Diverse Shapes from Natural Language.** 18339-18348

Shang Chai, Liansheng Zhuang, Fengying Yan

LayoutDM: Transformer-based Diffusion Model for Layout Generation. 18349-18358

Su Wang, Chitwan Saharia, Ceslee Montgomery, Jordi Pont-Tuset, Shai Noy, Stefano Pellegrini, Yasumasa Onoe, Sarah Laszlo, David J. Fleet, Radu Soricut, Jason Baldridge, Mohammad Norouzi, Peter Anderson, William Chan:

Imagen Editor and EditBench: Advancing and Evaluating Text-Guided Image Inpainting. 18359-18369Omri Avrahami, Thomas Hayes, Oran Gafni, Sonal Gupta, Yaniv Taigman, Devi Parikh, Dani Lischinski, Ohad Fried, Xi Yin: **SpaText: Spatio-Textual Representation for Controllable Image Generation.** 18370-18380Binxin Yang, Shuyang Gu, Bo Zhang, Ting Zhang, Xuejin Chen, Xiaoyan Sun, Dong Chen, Fang Wen: **Paint by Example: Exemplar-based Image Editing with Diffusion Models.** 18381-18391Tim Brooks, Aleksander Holynski, Alexei A. Efros: **InstructPix2Pix: Learning to Follow Image Editing Instructions.** 18392-18402Zhaoyun Jiang, Jiaqi Guo, Shizhao Sun, Huayu Deng, Zhongkai Wu, Vuksan Mijovic, Zijiang James Yang, Jian-Guang Lou, Dongmei Zhang: **LayoutFormer++: Conditional Graphic Layout Generation via Constraint Serialization and Decoding Space Restriction.** 18403-18412Vincent Tao Hu, David W. Zhang, Yuki M. Asano, Gertjan J. Burghouts, Cees G. M. Snoek: **Self-Guided Diffusion Models.** 18413-18422Animesh Karnewar, Andrea Vedaldi, David Novotný, Niloy J. Mitra: **HOLODIFFUSION: Training a 3D Diffusion Model Using 2D Images.** 18423-18433Yiming Qin, Huangjie Zheng, Jiangchao Yao, Mingyuan Zhou, Ya Zhang : **Class-Balancing Diffusion Models.** 18434-18443Haomiao Ni, Changhao Shi, Kai Li, Sharon X. Huang, Martin Renqiang Min : **Conditional Image-to-Video Generation with Latent Flow Diffusion Models.** 18444-18455Sihyun Yu, Kihyuk Sohn, Subin Kim, Jinwoo Shin: **Video Probabilistic Diffusion Models in Projected Latent Space.** 18456-18466Jiahui Zhang, Fangneng Zhan, Christian Theobalt, Shijian Lu: **Regularized Vector Quantization for Tokenized Image Synthesis.** 18467-18476Lishun Wang, Miao Cao, Xin Yuan: **EfficientSCI: Densely Connected Network with Space-time Factorization for Large-scale Video Snapshot Compressive Imaging.** 18477-18486Bowen Liu, Yu Chen, Rakesh Chowdary Machineni, Shiyu Liu, Hun-Seok Kim: **MMVC: Learned Multi-Mode Video Compression with Block-based Prediction Mode Selection and Density-Adaptive Entropy Coding.** 18487-18496Carlos Gomes, Roberto Azevedo, Christopher Schroers: **Video Compression with Entropy-Constrained Neural Representations.** 18497-18506Vishwanath Saragadam, Daniel Lejeune, Jasper Tan, Guha Balakrishnan, Ashok Veeraraghavan, Richard G. Baraniuk: **WIRE: Wavelet Implicit Neural Representations.** 18507-18516Runzhao Yang: **TINC: Tree-Structured Implicit Neural Compression.** 18517-18526Youmin Zhang, Xianda Guo, Matteo Poggi, Zheng Zhu, Guan Huang, Stefano Mattoccia : **CompletionFormer: Depth Completion with Convolutions and Vision Transformers.** 18527-18536Ning Zhang, Francesco Nex , George Vosselman, Norman Kerle : **Lite-Mono: A Lightweight CNN and Transformer Architecture for Self-Supervised Monocular Depth Estimation.** 18537-18546Huanrui Yang, Hongxu Yin, Maying Shen, Pavlo Molchanov, Hai Li, Jan Kautz: **Global Vision Transformer Pruning with Hessian-Aware Saliency.** 18547-18557Feng Li, Ailing Zeng, Shilong Liu, Hao Zhang, Hongyang Li, Lei Zhang, Lionel M. Ni: **Lite DETR : An Interleaved Multi-Scale Encoder for Efficient DETR.** 18558-18567Ryan Grainger, Thomas Paniagua, Xi Song, Naresh Cuntoor, Mun Wai Lee, Tianfu Wu : **PaCa-ViT: Learning Patch-to-Cluster Attention in Vision Transformers.** 18568-18578Sora Takashima, Ryo Hayamizu, Nakamasa Inoue, Hirokatsu Kataoka, Rio Yokota: **Visual Atoms: Pre-Training Vision Transformers with Sinusoidal Waves.** 18579-18588Hao Lu, Zitong Yu, Xuesong Niu, Yingcong Chen: **Neuron Structure Modeling for Generalizable Remote Physiological Measurement.** 18589-18599Stefan Kolek, Robert Windesheim, Héctor Andrade-Loarca, Gitta Kutyniok, Ron Levie: **Explaining Image Classifiers with Multiscale Directional Image Representation.** 18600-18609Yunjie Tian, Lingxi Xie, Zhaozhi Wang, Longhui Wei, Xiaopeng Zhang, Jianbin Jiao, Yaowei Wang, Qi Tian, Qixiang Ye: **Integrally Pre-Trained Transformer Pyramid Networks.** 18610-18620Minsu Kim, Seungryong Kim, Jungin Park, Seongheon Park, Kwanghoon Sohn: **PartMix: Regularization Strategy to Learn Part Discovery for Visible-Infrared Person Re-Identification.** 18621-



18632

- Shuxuan Guo, Yinlin Hu, José M. Álvarez, Mathieu Salzmann:
Knowledge Distillation for 6D Pose Estimation by Aligning Distributions of Local Predictions. 18633-18642
- Qiaqiao Wei, Hui Zhang, Jun-Hai Yong:
Focused and Collaborative Feedback Integration for Interactive Image Segmentation. 18643-18652
- Jiang Liu, Hui Ding, Zhaowei Cai, Yuting Zhang, Ravi Kumar Satzoda, Vijay Mahadevan, R. Manmatha:
PolyFormer: Referring Image Segmentation as Sequential Polygon Generation. 18653-18663
- Deunsol Jung, Sanghyun Kim, Won Hwa Kim, Minsu Cho:
Devil's on the Edges: Selective Quad Attention for Scene Graph Generation. 18664-18674
- Jingkang Yang, Wenzuan Peng, Xiangtai Li, Zujin Guo, Liangyu Chen, Bo Li, Zheng Ma, Kaiyang Zhou, Wayne Zhang , Chen Change Loy, Ziwei Liu:
Panoptic Video Scene Graph Generation. 18675-18685
- Shenyuan Gao, Chunluan Zhou, Jun Zhang:
Generalized Relation Modeling for Transformer Tracking. 18686-18695
- Haojie Zhao, Dong Wang, Huchuan Lu:
Representation Learning for Visual Object Tracking by Masked Appearance Transfer. 18696-18705
- Liulei Li, Wenguan Wang , Tianfei Zhou, Jianwu Li, Yi Yang:
Unified Mask Embedding and Correspondence Learning for Self-Supervised Video Segmentation. 18706-18716
- Ashish Singh, Michael J. Jones, Erik G. Learned-Miller:
EVAL: Explainable Video Anomaly Localization. 18717-18726
- Mingzhen Sun, Weining Wang, Xinxin Zhu, Jing Liu:
MOSO: Decomposing MOTion, Scene and Object for Video Prediction. 18727-18737
- Ali Athar, Alexander Hermans, Jonathon Luiten, Deva Ramanan, Bastian Leibe:
TarViS: A Unified Approach for Target-Based Video Segmentation. 18738-18748
- Md Mohaiminul Islam , Mahmudul Hasan, Kishan Shamsundar Athrey, Tony Braskich, Gedas Bertasius:
Efficient Movie Scene Detection using State-Space Transformers. 18749-18758
- Harshayu Girase, Nakul Agarwal, Chiho Choi, Karttikeya Mangalam:
Latency Matters: Real-Time Action Forecasting Transformer. 18759-18769
- Cheng Tan, Zhangyang Gao, Lirong Wu, Yongjie Xu, Jun Xia, Siyuan Li, Stan Z. Li:
Temporal Attention Unit: Towards Efficient Spatiotemporal Predictive Learning. 18770-18782
- Joanna Hong, Minsu Kim, Jeongsoo Choi, Yong Man Ro:
Watch or Listen: Robust Audio-Visual Speech Recognition with Visual Corruption Modeling and Reliability Scoring. 18783-18794
- Wei-Ning Hsu, Tal Remez, Bowen Shi, Jacob Donley, Yossi Adi:
ReViSE: Self-Supervised Speech Resynthesis with Visual Input for Universal and Generalized Speech Regeneration. 18796-18806
- Xubo Liu, Egor Lakomkin, Konstantinos Vougioukas, Pingchuan Ma, Honglie Chen, Ruiming Xie, Morrie Doulaty, Niko Moritz, Jáchym Kolár, Stavros Petridis, Maja Pantic, Christian Fuegen:
SynthVSR: Scaling Up Visual Speech RecognitionWith Synthetic Supervision. 18806-18815
- Zhen Xing, Qi Dai, Han Hu, Jingjing Chen , Zuxuan Wu, Yu-Gang Jiang:
SVFormer: Semi-supervised Video Transformer for Action Recognition. 18816-18826
- Junyu Gao, Mengyuan Chen, Changsheng Xu:
Collecting Cross-Modal Presence-Absence Evidence for Weakly-Supervised Audio- Visual Event Perception. 18827-18836
- Sauradip Nag, Xiatian Zhu, Yi-Zhe Song, Tao Xiang:
Post-Processing Temporal Action Detection. 18837-18845
- Anshul Shah, Aniket Roy, Ketul Shah, Shlok Mishra, David Jacobs, Anoop Cherian, Rama Chellappa:
HaLP: Hallucinating Latent Positives for Skeleton-based Self-Supervised Learning of Actions. 18846-18856
- Dingfeng Shi, Yujie Zhong, Qiong Cao, Lin Ma, Jia Li, Dacheng Tao:
TriDet: Temporal Action Detection with Relative Boundary Modeling. 18857-18866
- Aayush Jung Rana, Yogesh S. Rawat:
Hybrid Active Learning via Deep Clustering for Video Action Detection. 18867-18877
- Yu Wang, Yadong Li, Hongbin Wang:
Two-Stream Networks for Weakly-Supervised Temporal Action Localization with Semantic-Aware Mechanisms. 18878-18887
- Zhicheng Zhang , Lijuan Wang, Jufeng Yang:
Weakly Supervised Video Emotion Detection and Prediction via Cross-Modal Temporal Erasing Network. 18888-18897
- Bei Gan, Xiujun Shu, Ruizhi Qiao, Haoqian Wu, Keyu Chen, Hanjun Li, Bo Ren:
Collaborative Noisy Label Cleaner: Learning Scene-aware Trailers for Multi-modal Highlight Detection in Movies. 18898-18907



- Yifei Huang, Lijin Yang, Yoichi Sato:
Weakly Supervised Temporal Sentence Grounding with Uncertainty-Guided Self-training. 18908-18918
- Yi Li, Kyle Min, Subarna Tripathi, Nuno Vasconcelos:
SViT: Temporal Learning of Sparse Video-Text Transformers. 18919-18929
- Tengda Han, Max Bain, Arsha Nagrani, Gü̈l Varol, Weidi Xie, Andrew Zisserman:
AutoAD: Movie Description in Context. 18930-18940
- Xin Gu, Guang Chen , Yufei Wang, Libo Zhang, Tiejian Luo, Longyin Wen:
Text with Knowledge Graph Augmented Transformer for Video Captioning. 18941-18951
- Nikita Dvornik, Isma Hadji, Ran Zhang, Konstantinos G. Derpanis, Richard P. Wildes, Allan D. Jepson:
StepFormer: Self-Supervised Step Discovery and Localization in Instructional Videos. 18952-18961
- Xiaoshuai Hao, Wanqian Zhang, Dayan Wu, Fei Zhu, Bo Li:
Dual Alignment Unsupervised Domain Adaptation for Video-Text Retrieval. 18962-18972
- Chaolei Tan, Zihang Lin, Jian-Fang Hu, Wei-Shi Zheng, Jianhuang Lai:
Hierarchical Semantic Correspondence Networks for Video Paragraph Grounding. 18973-18982
- Renjing Pei, Jianzhuang Liu, Weimian Li, Bin Shao, Songcen Xu, Peng Dai, Juwei Lu, Youliang Yan:
CLIPPING: Distilling CLIP-Based Models with a Student Base for Video-Language Retrieval. 18983-18992
- Sitao Zhang, Yimu Pan, James Z. Wang:
Learning Emotion Representations from Verbal and Nonverbal Communication. 18993-19004
- Dingkang Yang, Zhaoyu Chen, Yuzheng Wang, Shunli Wang, Mingcheng Li, Siao Liu, Xiao Zhao, Shuai Huang, Zhiyan Dong, Peng Zhai, Lihua Zhang:
Context De-Confounded Emotion Recognition. 19005-19015
- Yiting Cheng, Fangyun Wei, Jianmin Bao, Dong Chen, Wenqiang Zhang:
CiCo: Domain-Aware Sign Language Retrieval via Cross-Lingual Contrastive Learning. 19016-19026
- Chuanqi Zang, Hanqing Wang, Mingtao Pei, Wei Liang:
Discovering the Real Association: Multimodal Causal Reasoning in Video Question Answering. 19027-19036
- QiuHong Anna Wei, Sijie Ding, Jeong Joon Park , Rahul Sajnani, Adrien Poulenard, Srinath Sridhar, Leonidas J. Guibas:
LEGO-Net: Learning Regular Rearrangements of Objects in Rooms. 19037-19047
- Xiaohan Wang, Wenguan Wang , Jiayi Shao, Yi Yang:
LANA: A Language-Capable Navigator for Instruction Following and Generation. 19048-19058
- Yuying Ge, Annabella Macaluso, Li Erran Li, Ping Luo, Xiaolong Wang:
Policy Adaptation from Foundation Model Feedback. 19059-19069
- Michael S. Ryoo, Keerthana Gopalakrishnan, Kumara Kahatapitiya, Ted Xiao, Kanishka Rao, Austin Stone, Yao Lu, Julian Ibarz, Anurag Arnab:
Token Turing Machines. 19070-19081
- Steven Spratley, Krista A. Ehinger, Tim Miller:
Unicode Analogies: An Anti-Objectivist Visual Reasoning Challenge. 19082-19091
- Chuanhao Li, Zhen Li, Chenchen Jing, Yunde Jia, Yuwei Wu:
Exploring the Effect of Primitives for Compositional Generalization in Vision-and-Language. 19092-19101
- Xi Zhang, Feifei Zhang, Changsheng Xu:
VQACL: A Novel Visual Question Answering Continual Learning Setting. 19102-19112
- Muhammad Uzair Khattak, Hanoona Abdul Rasheed, Muhammad Maaz, Salman H. Khan, Fahad Shahbaz Khan:
MaPLe: Multi-modal Prompt Learning. 19113-19122
- Chun-Hsiao Yeh, Bryan C. Russell, Josef Sivic, Fabian Caba Heilbron, Simon Jenni:
Meta-Personalizing Vision-Language Models to Find Named Instances in Video. 19123-19132
- Aochuan Chen, Yuguang Yao, Pin-Yu Chen, Yihua Zhang, Sijia Liu:
Understanding and Improving Visual Prompting: A Label-Mapping Perspective. 19133-19143
- Jiamu Sun, Gen Luo, Yiyi Zhou, Xiaoshuai Sun, Guannan Jiang, Zhiyu Wang, Rongrong Ji:
RefTeacher: A Strong Baseline for Semi-Supervised Referring Expression Comprehension. 19144-19154
- Yunhao Gou, Tom Ko, Hansi Yang, James T. Kwok, Yu Zhang, Mingxuan Wang:
Leveraging per Image-Token Consistency for Vision-Language Pre-training. 19155-19164
- Ziyan Yang, Kushal Kafle, Franck Dernoncourt, Vicente Ordonez:
Improving Visual Grounding by Encouraging Consistent Gradient-Based Explanations. 19165-19174
- Wenhui Wang, Hangbo Bao, Li Dong, Johan Bjorck, Zhiliang Peng, Qiang Liu, Kriti Aggarwal, Owais Khan Mohammed, Saksham Singhal, Subhajit Som, Furu Wei:
Image as a Foreign Language: BEIT Pretraining for Vision and Vision-Language Tasks. 19175-19186
- Yue Yang, Artemis Panagopoulou, Shenghao Zhou, Daniel Jin, Chris Callison-Burch, Mark Yatskar:
Language in a Bottle: Language Model Guided Concept Bottlenecks for Interpretable Image Classification. 19187-19197
- Jinwoo Kim, Janghyuk Choi, Ho-Jin Choi, Seon Joo Kim:
Shepherding Slots to Objects: Towards Stable and Robust Object-Centric Learning. 19198-19207



- Mohamed El Banani, Karan Desai, Justin Johnson:
Learning Visual Representations via Language-Guided Sampling. 19208-19220
- Zheng Chang, Shuchen Weng, Peixuan Zhang, Yu Li , Si Li, Boxin Shi:
L-ColIns: Language-based Colorization With Instance Awareness. 19221-19230
- Yanmin Wu, Xinhua Cheng, Renrui Zhang, Zesen Cheng, Jian Zhang:
EDA: Explicit Text-Decoupling and Dense Alignment for 3D Visual Grounding. 19231-19242
- Jianyang Gu, Kai Wang, Hao Luo, Chen Chen, Wei Jiang, Yuqiang Fang, Shanghang Zhang, Yang You, Jian Zhao:
MSINet: Twins Contrastive Search of Multi-Scale Interaction for Object ReID. 19243-19253
- Zineng Tang, Ziyi Yang, Guoxin Wang, Yuwei Fang, Yang Liu, Chenguang Zhu, Michael Zeng, Cha Zhang, Mohit Bansal:
Unifying Vision, Text, and Layout for Universal Document Processing. 19254-19264
- Chen-Wei Xie, Siyang Sun, Xiong Xiong , Yun Zheng, Deli Zhao, Jingren Zhou:
RA-CLIP: Retrieval Augmented Contrastive Language-Image Pre-Training. 19265-19274
- Zhengxin Pan, Fangyu Wu, Bailing Zhang:
Fine-grained Image-text Matching by Cross-modal Hard Aligning Network. 19275-19284
- Xiwen Wei, Zhen Xu, Cheng Liu, Si Wu, Zhiwen Yu, Hau-San Wong:
Text-Guided Unsupervised Latent Transformation for Multi-Attribute Image Manipulation. 19285-19294
- Ahmet Iscen, Alireza Fathi, Cordelia Schmid:
Improving Image Recognition by Retrieving from Web-Scale Image-Text Data. 19295-19304
- Kuniaki Saito, Kihyuk Sohn, Xiang Zhang, Chun-Liang Li, Chen-Yu Lee, Kate Saenko, Tomas Pfister:
Pic2Word: Mapping Pictures to Words for Zero-shot Composed Image Retrieval. 19305-19314
- Haoyuan Li, Hao Jiang, Tao Jin, Mengyan Li, Yan Chen, Zhijie Lin, Yang Zhao, Zhou Zhao:
DATE: Domain Adaptive Product Seeker for E-Commerce. 19315-19324
- Zhiqiu Lin, Samuel Yu , Zhiyi Kuang , Deepak Pathak, Deva Ramanan:
Multimodality Helps Unimodality: Cross-Modal Few-Shot Learning with Multimodal Models. 19325-19337
- Sachin Goyal, Ananya Kumar, Sankalp Garg, Zico Kolter, Aditi Raghunathan:
Finetune like you pretrain: Improved finetuning of zero-shot vision models. 19338-19347
- Maoyuan Ye, Jing Zhang, Shanshan Zhao, Juhua Liu, Tongliang Liu, Bo Du, Dacheng Tao:
DeepSolo: Let Transformer Decoder with Explicit Points Solo for Text Spotting. 19348-19357
- Yuxin Fang, Wen Wang, Binhui Xie, Quan Sun, Ledell Wu, Xinggang Wang, Tiejun Huang, Xinlong Wang, Yue Cao:
EVA: Exploring the Limits of Masked Visual Representation Learning at Scale. 19358-19369
- Sijie Zhu, Linjie Yang, Chen Chen, Mubarak Shah, Xiaohui Shen, Heng Wang:
\$R^2\$ Former: Unified Retrieval and Reranking Transformer for Place Recognition. 19370-19380
- Shijie Wang, Jianlong Chang, Haojie Li, Zhihui Wang, Wanli Ouyang, Qi Tian:
Open-Set Fine-Grained Retrieval via Prompting Vision-Language Evaluator. 19381-19391
- Sipeng Zheng, Boshen Xu , Qin Jin:
Open-Category Human-Object Interaction Pre-training via Language Modeling Framework. 19392-19402
- Dolev Ofri-Amar, Michal Geyer, Yoni Kasten, Tali Dekel:
Neural Congealing: Aligning Images to a Joint Semantic Atlas. 19403-19412
- Jishnu Mukhoti, Tsung-Yu Lin, Omid Poursaeed, Rui Wang, Ashish Shah, Philip H. S. Torr, Ser-Nam Lim:
Open Vocabulary Semantic Segmentation with Patch Aligned Contrastive Learning. 19413-19423
- Jie Yang, Chaoqun Wang, Zhen Li, Junle Wang, Ruimao Zhang :
Semantic Human Parsing via Scalable Semantic Transfer Over Multiple Label Domains. 19424-19433
- Weihuang Liu , Xi Shen, Chi-Man Pun, Xiaodong Cun:
Explicit Visual Prompting for Low-Level Structure Segmentations. 19434-19445
- Jie Qin, Jie Wu, Pengxiang Yan, Ming Li, Yuxi Ren, Xuefeng Xiao, Yitong Wang, Rui Wang, Shilei Wen, Xin Pan, Xingang Wang:
FreeSeg: Unified, Universal and Open-Vocabulary Image Segmentation. 19446-19455
- Seonghoon Yu , Paul Hongsuck Seo, Jeany Son:
Zero-shot Referring Image Segmentation with Global-Local Context Features. 19456-19465
- Shubhankar Borse, Debasmit Das, Hyojin Park, Hong Cai, Risheek Garrepalli, Fatih Porikli:
DejaVu: Conditional Regenerative Learning to Enhance Dense Prediction. 19466-19477
- Li Xu, Mark He Huang, Xindi Shang, Zehuan Yuan, Ying Sun, Jun Liu:
Meta Compositional Referring Expression Segmentation. 19478-19487
- Minghao Zhou, Hong Wang, Qian Zhao, Yuexiang Li, Yawen Huang, Deyu Meng, Yefeng Zheng:
Interactive Segmentation as Gaussian Process Classification. 19488-19497
- Shuting He, Henghui Ding , Wei Jiang:
Semantic-Promoted Debiasing and Background Disambiguation for Zero-Shot Instance Segmentation. 19498-19507
- Tobias Kalb , Jürgen Beyerer:
Principles of Forgetting in Domain-Incremental Semantic Segmentation in Adverse Weather Conditions. 19508-19518



- Mingxiang Liao, Zonghao Guo, Yuze Wang, Peng Yuan, Bailan Feng, Fang Wan:
AttentionShift: Iteratively Estimated Part-Based Attention Map for Pointly Supervised Instance Segmentation. 19519-19528
- Jiacong Xu, Zixiang Xiong, Shankar P. Bhattacharyya:
PIDNet: A Real-time Semantic Segmentation Network Inspired by PID Controllers. 19529-19539
- Hyun Seok Seong, WonJun Moon, Su Been Lee, Jae-Pil Heo:
Leveraging Hidden Positives for Unsupervised Semantic Segmentation. 19540-19549
- Zhisheng Zhong, Jiequan Cui, Yibo Yang, Xiaoyang Wu, Xiaojuan Qi, Xiangyu Zhang, Jiaya Jia:
Understanding Imbalanced Semantic Segmentation Through Neural Collapse. 19550-19559
- Yuchao Wang, Jingjing Fei, Haochen Wang, Wei Li, Tianpeng Bao, Liwei Wu, Rui Zhao, Yujun Shen:
Balancing Logit Variation for Long-Tailed Semantic Segmentation. 19561-19573
- Shenghai Rong, Bohai Tu, Zilei Wang, Junjie Li:
Boundary-enhanced Co-training for Weakly Supervised Semantic Segmentation. 19574-19584
- Zicheng Wang, Zhen Zhao , Xiaoxia Xing, Dong Xu, Xiangyu Kong, Luping Zhou:
Conflict-Based Cross-View Consistency for Semi-Supervised Semantic Segmentation. 19585-19595
- Lian Xu, Wanli Ouyang, Mohammed Bennamoun , Farid Boussaïd, Dan Xu:
Learning Multi-Modal Class-Specific Tokens for Weakly Supervised Dense Object Localization. 19596-19605
- Jongheon Jeong, Yang Zou, Taewan Kim, Dongqing Zhang, Avinash Ravichandran, Onkar Dabeer:
WinCLIP: Zero-/Few-Shot Anomaly Classification and Segmentation. 19606-19616
- Huayu Mai, Rui Sun, Tianzhu Zhang, Zhiwei Xiong, Feng Wu:
DualRel: Semi-Supervised Mitochondria Segmentation from A Prototype Perspective. 19617-19626
- Dahyun Kang, Piotr Koniusz, Minsu Cho, Naila Murray:
Distilling Self-Supervised Vision Transformers for Weakly-Supervised Few-Shot Classification & Segmentation. 19627-19638
- Yang Wu, Huihui Song, Bo Liu, Kaihua Zhang, Dong Liu:
Co-Salient Object Detection with Uncertainty-Aware Group Exchange-Masking. 19639-19648
- Han Lin, Guangxing Han, Jiawei Ma, Shiyuan Huang, Xudong Lin, Shih-Fu Chang:
Supervised Masked Knowledge Distillation for Few-Shot Transformers. 19649-19659
- Xinyu Tian, Jing Zhang, Mochu Xiang, Yuchao Dai:
Modeling the Distributional Uncertainty for Salient Object Detection Models. 19660-19670
- Xuanyi Du, Weitao Wan, Chong Sun, Chen Li:
Weak-shot Object Detection through Mutual Knowledge Transfer. 19671-19680
- Shuailei Ma, Yuefeng Wang, Ying Wei, Jiaqi Fan, Thomas H. Li, Hongli Liu, Fanbing Lv:
CAT: LoCalization and IdentificAtion Cascade Detection Transformer for Open-World Object Detection. 19681-19690
- Xiao Zhou, Yujie Zhong, Zhen Cheng, Fan Liang, Lin Ma:
Adaptive Sparse Pairwise Loss for Object Re-Identification. 19691-19701
- Ding Jia, Yuhui Yuan, Haodi He, Xiaopei Wu, Haojun Yu, Weihong Lin, Lei Sun, Chao Zhang, Han Hu:
DETRs with Hybrid Matching. 19702-19712
- Jingyi Xu, Hieu Le, Dimitris Samaras:
Generating Features with Increased Crop-Related Diversity for Few-Shot Object Detection. 19713-19722
- Yichen Zhu, Qiqi Zhou, Ning Liu, Zhiyuan Xu, Zhicai Ou, Xiaofeng Mou, Jian Tang:
ScaleKD: Distilling Scale-Aware Knowledge in Small Object Detector. 19723-19733
- Bimsara Pathiraja, Malitha Gunawardhana , Muhammad Haris Khan:
Multiclass Confidence and Localization Calibration for Object Detection. 19734-19743
- Geeho Kim, Junoh Kang, Bohyung Han:
Open-Set Representation Learning through Combinatorial Embedding. 19744-19753
- Tianyi Ma, Yifan Sun, Zongxin Yang, Yi Yang:
ProD: Prompting-to-disentangle Domain Knowledge for Cross-domain Few-shot Image Classification. 19754-19763
- Ming Y. Lu, Bowen Chen, Andrew Zhang, Drew F. K. Williamson, Richard J. Chen, Tong Ding , Long Phi Le, Yung-Sung Chuang, Faisal Mahmood:
Visual Language Pretrained Multiple Instance Zero-Shot Transfer for Histopathology Images. 19764-19775
- Weijie Chen, Xinyan Wang, Yuhang Wang:
FFF: Fragment-Guided Flexible Fitting for Building Complete Protein Structures. 19776-19785
- Hritam Basak, Zhaozheng Yin :
Pseudo-Label Guided Contrastive Learning for Semi-Supervised Medical Image Segmentation. 19786-19797
- Cheng Jiang, Xinhai Hou, Akhil Kondepudi , Asadur Chowdury , Christian W. Freudiger, Daniel A. Orringer , Honglak Lee, Todd C. Hollon :
Hierarchical Discriminative Learning Improves Visual Representations of Biomedical Microscopy. 19798-19808
- Zhongzhen Huang, Xiaofan Zhang, Shaoting Zhang:
KiUT: Knowledge-injected U-Transformer for Radiology Report Generation. 19809-19818



- Haoxuan Che, Siyu Chen, Hao Chen:
Image Quality-aware Diagnosis via Meta-knowledge Co-embedding. 19819-19829
- Tiancheng Lin, Zhimiao Yu, Hongyu Hu, Yi Xu, Chang Wen Chen:
Interventional Bag Multi-Instance Learning On Whole-Slide Pathological Images. 19830-19839
- Kihyuk Sohn, Huiwen Chang, José Lezama, Luisa Polania, Han Zhang, Yuan Hao, Irfan Essa, Lu Jiang:
Visual Prompt Tuning for Generative Transfer Learning. 19840-19851
- Yong Hyun Ahn, Gyeong-Moon Park, Seong Tae Kim:
LINe: Out-of-Distribution Detection by Leveraging Important Neurons. 19852-19862
- Weiqing Yan, Yuanyang Zhang, Chenlei Lv, Chang Tang, Guanghui Yue, Liang Liao, Weisi Lin:
GCFAgg: Global and Cross-View Feature Aggregation for Multi-View Clustering. 19863-19872
- Mengyao Xie, Zongbo Han, Changqing Zhang, Yichen Bai, Qinghua Hu:
Exploring and Exploiting Uncertainty for Incomplete Multi-View Classification. 19873-19882
- Shuo Yang, Zhaopan Xu, Kai Wang, Yang You, Hongxun Yao, Tongliang Liu, Min Xu:
BiCro: Noisy Correspondence Rectification for Multi-modality Data via Bi-directional Cross-modal Similarity Consistency. 19883-19892
- Zhicai Wang, Yanbin Hao, Tingting Mu, Ouxiang Li, Shuo Wang, Xiangnan He:
Bi-Directional Distribution Alignment for Transductive Zero-Shot Learning. 19893-19902
- Sungyeon Kim, Boseung Jeong, Suha Kwak:
HIER: Metric Learning Beyond Class Labels via Hierarchical Regularization. 19903-19912
- Chen Feng, Ioannis Patras:
MaskCon: Masked Contrastive Learning for Coarse-Labelled Dataset. 19913-19922
- Rohit Gupta, Anirban Roy, Claire Christensen, Sujeong Kim, Sarah Gerard, Madeline Cincebeaux, Ajay Divakaran, Todd Grindal, Mubarak Shah:
Class Prototypes based Contrastive Learning for Classifying Multi-Label and Fine-Grained Educational Videos. 19923-19933
- Yuanpeng Tu, Boshen Zhang, Yuxi Li, Liang Liu, Jian Li, Yabiao Wang, Chengjie Wang, Cairong Zhao:
Learning from Noisy Labels with Decoupled Meta Label Purifier. 19934-19943
- Yingjun Du, Jiayi Shen, Xiantong Zhen, Cees G. M. Snoek:
SuperDisco: Super-Class Discovery Improves Visual Recognition for the Long-Tail. 19944-19954
- Matthias Eisenmann, Annika Reinke, Vivienn Weru, Minu Dietlinde Tizabi, Fabian Isensee, Tim J. Adler, Sharib Ali , Vincent Andreczyk, Marc Aubreville, Ujjwal Baid, Spyridon Bakas, Niranjan Balu, Sophia Bano , Jorge Bernal, Sebastian Bodenstedt, Alessandro Casella , Veronika Cheplygina, Marie Daum, Marleen de Bruijne, Adrien Depeursinge, Reuben Dorent , Jan Egger, David G. Ellis, Sandy Engelhardt, Melanie Ganz, Noha M. Ghatwary, Gabriel Girard, Patrick Godau, Anubha Gupta, Lasse Hansen, Kanako Harada, Mattias P. Heinrich, Nicholas Heller , Alessa Hering, Arnaud Huaulmé, Pierre Jannin , A. Emre Kavur, Oldrich Kodym, Michal Kozubek , Jianning Li, Hongwei Bran Li, Jun Ma, Carlos Martín-Isla, Bjoern H. Menze, J. Alison Noble, Valentin Oreiller, Nicolas Padoy, Sarthak Pati, Kelly Payette, Tim Rädsch, Jonathan Rafael-Patino, Vivek Singh Bawa, Stefanie Speidel , Carole H. Sudre , Kimberlin M. H. van Wijnen, Martin Wagner , D. Wei, Amine Yamlahi, Moi Hoon Yap, C. Yuan, Maximilian Zenk, A. Zia, David Zimmerer, Dogu Baran Aydogan, Binod Bhattacharai , Louise Bloch , Raphael Brügel , J. Cho, C. Choi, Q. Dou, Ivan Ezhov, Christoph M. Friedrich , C. Fuller, Rebati Raman Gaire, Adrian Galdran, Álvaro García-Faura, Maria Grammatikopoulou, S. Hong, Mostafa Jahanifar, I. Jang, Abdolrahim Kadkhodamohammadi, I. Kang, Florian Kofler, S. Kondo, Hugo Jaco Kuijf, M. Li, M. Luu, Tomaz Martincic, Pedro Morais , Mohamed A. Naser , Bruno Oliveira, David Owen , S. Pang, J. Park, S. Park, Szymon Plotka, Élodie Puybareau, Nasir M. Rajpoot, K. Ryu, Numan Saeed, Adam Shephard, Pengcheng Shi, Dejan Stepec, Ronast Subedi, Guillaume Tochon, Helena R. Torres, Hélène Urien, João L. Vilaça , Kareem A. Wahid , H. Wang, J. Wang, L. Wang, X. Wang, Benedikt Wiestler, Marek Wodzinski , F. Xia, J. Xie, Z. Xiong, Sen Yang , Y. Yang, Z. Zhao, Klaus H. Maier-Hein, Paul F. Jäger, Annette Kopp-Schneider, Lena Maier-Hein:
Why is the Winner the Best? 19955-19966
- Emanuel Sanchez Aimar, Arvi Jonnarth , Michael Felsberg, Marco Kuhlmann:
Balanced Product of Calibrated Experts for Long-Tailed Recognition. 19967-19977
- Jiahao Chen, Bing Su:
Transfer Knowledge from Head to Tail: Uncertainty Calibration under Long-tailed Distribution. 19978-19987
- Thanh-Dat Truong, Ngan Le , Bhiksha Raj, Jackson David Cothren , Khoa Luu:
FREDOM: Fairness Domain Adaptation Approach to Semantic Scene Understanding. 19988-19997
- Yang Liu, Zhipeng Zhou, Baigui Sun:
COT: Unsupervised Domain Adaptation with Clustering and Optimal Transport. 19998-20007
- Fan Wang, Zhongyi Han, Zhiyan Zhang, Rundong He, Yilong Yin:
MHPL: Minimum Happy Points Learning for Active Source Free Domain Adaptation. 20008-20018
- Sanqing Qu, Tianpei Zou, Florian Röhrbein, Cewu Lu, Guang Chen, Dacheng Tao, Changjun Jiang:
Upcycling Models Under Domain and Category Shift. 20019-20028
- Yunfeng Fan, Wenchao Xu, Haozhao Wang, Junxiao Wang , Song Guo:
PMR: Prototypical Modal Rebalance for Multimodal Learning. 20029-20038
- Shicai Wei, Chunbo Luo, Yang Luo:
MMANet: Margin-Aware Distillation and Modality-Aware Regularization for Incomplete Multimodal Learning.



- 20039-20049
Shuai Wang , Daoan Zhang , Zipei Yan , Jianguo Zhang, Rui Li:
Feature Alignment and Uniformity for Test Time Adaptation. 20050-20060
- 20040-20049
Fei Zhou, Peng Wang, Lei Zhang, Wei Wei, Yanning Zhang:
Revisiting Prototypical Network for Cross Domain Few-Shot Learning. 20061-20070
- 20049-20058
Zhiheng Li, Ivan Evtimov, Albert Gordo, Caner Hazirbas, Tal Hassner, Cristian Canton-Ferrer, Chenliang Xu, Mark Ibrahim:
A Whac-A-Mole Dilemma: Shortcuts Come in Multiples Where Mitigating One Amplifies Others. 20071-20082
- 20058-20067
Dmitry Senushkin, Nikolay Patakin, Arseny Kuznetsov, Anton Konushin:
Independent Component Alignment for Multi-Task Learning. 20083-20093
- 20067-20076
Shiguang Wang, Tao Xie, Jian Cheng, Xingcheng Zhang, Haijun Liu:
MDL-NAS: A Joint Multi-domain Learning Framework for Vision Transformer. 20094-20104
- 20076-20085
Dohwan Ko, Joonmyung Choi, Hyeong Kyu Choi, Kyoung-Woon On, Byungseok Roh, Hyunwoo J. Kim:
MELTR: Meta Loss Transformer for Learning to Fine-tune Video Foundation Models. 20105-20115
- 20085-20094
Dongshuo Yin , Yiran Yang, Zhechao Wang, Hongfeng Yu, Kaiwen Wei, Xian Sun:
1% VS 100%: Parameter-Efficient Low Rank Adapter for Dense Predictions. 20116-20126
- 20094-20103
Sungmin Cha, Sungjun Cho, Dasol Hwang, Sunwon Hong, Moontae Lee, Taesup Moon:
Rebalancing Batch Normalization for Exemplar-Based Class-Incremental Learning. 20127-20136
- 20103-20112
Jingwen Ye, Songhua Liu, Xinchao Wang:
Partial Network Cloning. 20137-20146
- 20112-20121
Shen Lin , Xiaoyu Zhang, Chenyang Chen, Xiaofeng Chen, Willy Susilo :
ERM-KTP: Knowledge-Level Machine Unlearning via Knowledge Transfer. 20147-20155
- 20121-20130
Jingzhi Li , Zidong Guo, Hui Li, Seungju Han , Ji-Won Baek, Min Yang, Ran Yang, Sungjoo Suh:
Rethinking Feature-based Knowledge Distillation for Face Recognition. 20156-20165
- 20130-20139
Zhicheng Sun, Yadong Mu, Gang Hua:
Regularizing Second-Order Influences for Continual Learning. 20166-20175
- 20139-20148
Tianli Zhang, Mengqi Xue, Jiangtao Zhang, Haofei Zhang , Yu Wang, Lechao Cheng , Jie Song, Mingli Song:
Generalization Matters: Loss Minima Flattening via Parameter Hybridization for Efficient Online Knowledge Distillation. 20176-20185
- 20148-20157
Wenju Sun , Qingyong Li, Jing Zhang, Wen Wang, Yangli-ao Geng:
Decoupling Learning and Remembering: a Bilevel Memory Framework with Knowledge Projection for Task-Incremental Learning. 20186-20195
- 20157-20166
Dongwan Kim, Bohyung Han:
On the Stability-Plasticity Dilemma of Class-Incremental Learning. 20196-20204
- 20166-20175
AmirMohammad Sarfi, Zahra Karimpour, Muawiz Chaudhary, Nasir Mohammad Khalid, Mirco Ravanelli, Sudhir Mudur, Eugene Belilovsky:
Simulated Annealing in Early Layers Leads to Better Generalization. 20205-20214
- 20175-20184
Qiang He, Huangyuan Su, Jieyu Zhang, Xinwen Hou:
Frustratingly Easy Regularization on Representation Can Boost Deep Reinforcement Learning. 20215-20225
- 20184-20193
Matteo Maggioni, Thomas Tanay, Francesca Babiloni, Steven McDonagh , Ales Leonardis:
Tunable Convolutions with Parametric Multi-Loss Optimization. 20226-20236
- 20193-20202
Fidel A. Guerrero-Peña, Heitor Rapela Medeiros, Thomas Dubail, Masih Aminbeidokhti, Eric Granger, Marco Pedersoli:
Re-basin via implicit Sinkhorn differentiation. 20237-20246
- 20202-20211
Xingxuan Zhang, Renzhe Xu, Han Yu, Hao Zou, Peng Cui:
Gradient Norm Aware Minimization Seeks First-Order Flatness and Improves Generalization. 20247-20257
- 20211-20220
Mengqiao Han, Liyuan Pan, Xiabi Liu:
AstroNet: When Astrocyte Meets Artificial Neural Network. 20258-20268
- 20220-20229
Ning Ding, Yehui Tang, Kai Han, Chao Xu, Yunhe Wang:
Network Expansion For Practical Training Acceleration. 20269-20279
- 20229-20238
Jie Ren, Mingjie Li, Qirui Chen, Huiqi Deng, Quanshi Zhang:
Defining and Quantifying the Emergence of Sparse Concepts in DNNs. 20280-20289
- 20238-20247
Isha Garg, Kaushik Roy:
Samples with Low Loss Curvature Improve Data Efficiency. 20290-20300
- 20247-20256
Yao Xiao, Ziyi Tang, Pengxu Wei, Cong Liu, Liang Lin:
Masked Images Are Counterfactual Samples for Robust Fine-Tuning. 20301-20310
- 20256-20265
Maan Qraitem, Kate Saenko, Bryan A. Plummer:
Bias Mimicking: A Simple Sampling Approach for Bias Mitigation. 20311-20320
- 20265-20274
Yijiang Liu , Huanrui Yang, Zhen Dong, Kurt Keutzer, Li Du, Shanghang Zhang:
NoisyQuant: Noisy Bias-Enhanced Post-Training Activation Quantization for Vision Transformers. 20321-20330
- 20274-20283
Guo-Hua Wang, Jianxin Wu:
Practical Network Acceleration with Tiny Sets. 20331-20340



- Devavrat Tomar , Guillaume Vray, Behzad Bozorgtabar, Jean-Philippe Thiran:
TeSLA: Test-Time Self-Learning With Automatic Adversarial Augmentation. 20341-20350
- Tie Hu, Mingbao Lin, Lizhou You, Fei Chao, Rongrong Ji:
Discriminator-Cooperated Feature Map Distillation for GAN Compression. 20351-20360
- Chen Chen, Daochang Liu, Siqi Ma, Surya Nepal , Chang Xu:
Private Image Generation with Dual-Purpose Auxiliary Classifier. 20361-20370
- Xiaodan Li, Yuefeng Chen, Yao Zhu, Shuhui Wang, Rong Zhang, Hui Xue:
ImageNet-E: Benchmarking Neural Network Robustness via Attribute Editing. 20371-20381
- Bin Ren, Yahui Liu, Yue Song, Wei Bi, Rita Cucchiara, Nicu Sebe , Wei Wang:
Masked Jigsaw Puzzle: A Versatile Position Embedding for Vision Transformers. 20382-20391
- Congqi Cao, Yue Lu, Peng Wang, Yanning Zhang:
A New Comprehensive Benchmark for Semi-supervised Video Anomaly Detection and Anticipation. 20392-20401
- Zhiqiang Liu, Yiming Zhou, Yuansheng Xu, Zilei Wang:
SimpleNet: A Simple Network for Image Anomaly Detection and Localization. 20402-20411
- Haozhao Wang, Yichen Li, Wenchao Xu, Ruixuan Li, Yufeng Zhan, Zhigang Zeng:
DaFKD: Domain-aware Federated Knowledge Distillation. 20412-20421
- Zixuan Qin, Liu Yang, Qilong Wang, Yahong Han, Qinghua Hu:
Reliable and Interpretable Personalized Federated Learning. 20422-20431
- Dongping Liao, Xitong Gao, Yiren Zhao, Chengzhong Xu:
Adaptive Channel Sparsity for Federated Learning under System Heterogeneity. 20432-20441
- Yuan-Yi Xu, Ci-Siang Lin, Yu-Chiang Frank Wang:
Bias-Eliminating Augmentation Learning for Debiased Federated Learning. 20442-20452
- Qianyu Zhou , Ke-Yue Zhang, Taiping Yao, Xuequan Lu , Ran Yi, Shouhong Ding, Lizhuang Ma:
Instance-Aware Domain Generalization for Face Anti-Spoofing. 20453-20463
- Guangrui Li, Guoliang Kang, Xiaohan Wang, Yunchao Wei, Yi Yang:
Adversarially Masking Synthetic to Mimic Real: Adaptive Noise Injection for Point Cloud Segmentation Adaptation. 20464-20474
- Liyanu Wang, Meng Wang, Daoqiang Zhang, Huazhu Fu :
Model Barrier: A Compact Un-Transferable Isolation Domain for Model Intellectual Property Protection. 20475-20484
- Qiuling Xu, Guanhong Tao, Jean Honorio , Yingqi Liu, Shengwei An, Guangyu Shen, Siyuan Cheng, Xiangyu Zhang:
MEDIC: Remove Model Backdoors via Importance Driven Cloning. 20485-20494
- Bingxu Mu, Zhenxing Niu, Le Wang, Xue Wang, Qiguang Miao, Rong Jin, Gang Hua:
Progressive Backdoor Erasing via connecting Backdoor and Adversarial Attacks. 20495-20503
- Gyojin Han, Jaehyun Choi, Haeil Lee, Junmo Kim:
Reinforcement Learning-Based Black-Box Model Inversion Attacks. 20504-20513
- Hao Huang, Ziyan Chen, Huanran Chen, Yongtao Wang, Kevin Zhang:
T-SEA: Transfer-Based Self-Ensemble Attack on Object Detection. 20514-20523
- Jérôme Rony, Jean-Christophe Pesquet, Ismail Ben Ayed:
Proximal Splitting Adversarial Attack for Semantic Segmentation. 20524-20533
- Zhibo Wang, Hongshan Yang, Yunhe Feng, Peng Sun, Hengchang Guo, Zhifei Zhang, Kui Ren:
Towards Transferable Targeted Adversarial Examples. 20534-20543
- Shenglin Yin, Kelu Yao, Sheng Shi, Yangzhou Du, Zhen Xiao :
AGAIN: Adversarial Training with Attribution Span Enlargement and Hybrid Feature Fusion. 20544-20553
- Hongjun Wang, Yisen Wang:
Generalist: Decoupling Natural and Robust Generalization. 20554-20563
- Yimu Wang, Dinghuai Zhang, Yihan Wu, Heng Huang, Hongyang Zhang:
Cooperation or Competition: Avoiding Player Domination for Multi-Target Robustness via Adaptive Budgets. 20564-20574
- Qian Li, Yuxiao Hu, Ye Liu, Dongxiao Zhang, Xin Jin, Yuntian Chen:
Discrete Point-Wise Attack is Not Enough: Generalized Manifold Adversarial Attack for Face Recognition. 20575-20584
- Han Liu , Yuhao Wu, Shixuan Zhai, Bo Yuan, Ning Zhang:
RIATIG: Reliable and Imperceptible Adversarial Text-to-Image Generation with Natural Prompts. 20585-20594
- Fahad Shamshad, Muzammal Naseer, Karthik Nandakumar:
CLIP2Protect: Protecting Facial Privacy Using Text-Guided Makeup via Adversarial Latent Search. 20595-20605
- Fabrizio Guillaro, Davide Cozzolino, Avneesh Sud, Nicholas Dufour, Luisa Verdoliva:
TruFor: Leveraging All-Round Clues for Trustworthy Image Forgery Detection and Localization. 20606-20615
- Jin Han, Yuta Asano , Boxin Shi, Yinjiang Zheng, Imari Sato:
High-fidelity Event-Radiance Recovery via Transient Event Frequency. 20616-20625



- Sara Sabour, Suhani Vora, Daniel Duckworth, Ivan Krasin, David J. Fleet, Andrea Tagliasacchi:
RobustNeRF: Ignoring Distractors with Robust Losses. 20626-20636
- Congyue Deng, Chiyu Max Jiang, Charles R. Qi, Xinchen Yan, Yin Zhou, Leonidas J. Guibas, Dragomir Anguelov:
NeRDi: Single-View NeRF Synthesis with Language-Guided Diffusion as General Image Priors. 20637-20647
- Jianchuan Chen, Wentao Yi, Liqian Ma, Xu Jia, Huchuan Lu:
GM-NeRF: Learning Generalizable Model-Based Neural Radiance Fields from Multi-View Images. 20648-20658
- Seunghyeon Seo, Donghoon Han, Yeonjin Chang, Nojun Kwak:
MixNeRF: Modeling a Ray with Mixture Density for Novel View Synthesis from Sparse Inputs. 20659-20668
- Ashkan Mirzaei, Tristan Aumentado-Armstrong, Konstantinos G. Derpanis, Jonathan Kelly, Marcus A. Brubaker, Igor Gilitschenski, Alex Levinstein:
SPIIn-NeRF: Multiview Segmentation and Perceptual Inpainting with Neural Radiance Fields. 20669-20679
- Daniel Rho, Byeonghyeon Lee, Seungtae Nam, Joo Chan Lee , Jong Hwan Ko, Eunbyung Park:
Masked Wavelet Representation for Compact Neural Radiance Fields. 20680-20690
- Zhengfei Kuang, Fujun Luan, Sai Bi, Zhixin Shu, Gordon Wetzstein, Kalyan Sunkavalli:
PaletteNeRF: Palette-based Appearance Editing of Neural Radiance Fields. 20691-20700
- Sicheng Li, Hao Li, Yue Wang, Yiyi Liao, Lu Yu:
SteerNeRF: Accelerating NeRF Rendering via Smooth Viewpoint Trajectory. 20701-20711
- Zicheng Zhang, Yinglu Liu, Congying Han, Yingwei Pan, Tiande Guo, Ting Yao:
Transforming Radiance Field with Lipschitz Network for Photorealistic 3D Scene Stylization. 20712-20721
- Chengxuan Zhu, Renjie Wan , Yunkai Tang, Boxin Shi:
Occlusion-Free Scene Recovery via Neural Radiance Fields. 20722-20731
- Tao Hu, Xiaogang Xu, Ruihang Chu, Jiaya Jia:
TriVol: Point Cloud Rendering via Triple Volumes. 20732-20741
- Ehsan Pajouheshgar, Yitao Xu, Tong Zhang, Sabine Süsstrunk:
DyNCA: Real-Time Dynamic Texture Synthesis Using Neural Cellular Automata. 20742-20751
- Haotong Lin, Qianqian Wang, Ruojin Cai, Sida Peng, Hadar Averbuch-Elor, Xiaowei Zhou, Noah Snavely:
Neural Scene Chronology. 20752-20761
- Marco Toschi, Riccardo De Matteo, Riccardo Spezialetti, Daniele De Gregorio, Luigi Di Stefano, Samuele Salti:
ReLight My NeRF: A Dataset for Novel View Synthesis and Relighting of Real World Objects. 20762-20772
- Kushagra Tiwary, Akshat Dave, Nikhil Behari, Tzofi Klinghoffer, Ashok Veeraraghavan, Ramesh Raskar:
ORCa: Glossy Objects as Radiance-Field Cameras. 20773-20782
- Yuekun Dai, Yihang Luo, Shangchen Zhou, Chongyi Li, Chen Change Loy:
Nighttime Smartphone Reflective Flare Removal Using Optical Center Symmetry Prior. 20783-20791
- Yifan Wang, Aleksander Holynski, Xiuming Zhang, Xuaner Zhang:
SunStage: Portrait Reconstruction and Relighting Using the Sun as a Light Stage. 20792-20802
- Geoffroi Côté, Fahim Mannan, Simon Thibault, Jean-François Lalonde, Felix Heide:
The Differentiable Lens: Compound Lens Search over Glass Surfaces and Materials for Object Detection. 20803-20812
- Ryo Kawahara, Meng-Yu Jennifer Kuo, Shohei Nobuhara:
Teleidoscopic Imaging System for Microscale 3D Shape Reconstruction. 20813-20822
- Jiaxiong Qiu , Peng-Tao Jiang, Yifan Zhu, Ze-Xin Yin, Ming-Ming Cheng, Bo Ren:
Looking Through the Glass: Neural Surface Reconstruction Against High Specular Reflections. 20823-20833
- Xiaoxiao Long, Cheng Lin, Lingjie Liu, Yuan Liu, Peng Wang, Christian Theobalt, Taku Komura, Wenping Wang:
NeuralUDF: Learning Unsigned Distance Fields for Multi-View Reconstruction of Surfaces with Arbitrary Topologies. 20834-20843
- Andreea Dogaru, Andrei-Timotei Ardelean, Savva Ignatyev, Egor Zakharov, Evgeny Burnaev:
Sphere-Guided Training of Neural Implicit Surfaces. 20844-20853
- Haim Sawdayee, Amir Vaxman, Amit H. Bermano:
OReX: Object Reconstruction from Planar Cross-sections Using Neural Fields. 20854-20862
- Lucy Chai, Richard Tucker, Zhengqi Li, Phillip Isola, Noah Snavely:
Persistent Nature: A Generative Model of Unbounded 3D Worlds. 20863-20874
- J. Ryan Shue, Eric Ryan Chan, Ryan Po, Zachary Ankner, Jiajun Wu, Gordon Wetzstein:
3D Neural Field Generation Using Triplane Diffusion. 20875-20886
- Jaehyeok Shim, Changwoo Kang, Kyungdon Joo:
Diffusion-Based Signed Distance Fields for 3D Shape Generation. 20887-20897
- Thomas Tanay, Ales Leonardis, Matteo Maggioni:
Efficient View Synthesis and 3D-based Multi-Frame Denoising with Multiplane Feature Representations. 20898-20907
- Jiale Xu, Xintao Wang, Weihao Cheng, Yan-Pei Cao, Ying Shan, Xiaohu Qie, Shenghua Gao:
Dream3D: Zero-Shot Text-to-3D Synthesis Using 3D Shape Prior and Text-to-Image Diffusion Models. 20908-20918



- Chong Bao, Yinda Zhang, Bangbang Yang, Tianxing Fan, Zesong Yang, Hujun Bao, Guofeng Zhang, Zhaopeng Cui:
SINE: Semantic-driven Image-based NeRF Editing with Prior-guided Editing Field. 20919-20929
- Dale Decatur, Itai Lang, Rana Hanocka:
3D Highlighter: Localizing Regions on 3D Shapes via Text Descriptions. 20930-20939
- Yushi Lan, Xuyi Meng, Shuai Yang, Chen Change Loy, Bo Dai:
Self-Supervised Geometry-Aware Encoder for Style-Based 3D GAN Inversion. 20940-20949
- Sizhe An, Hongyi Xu, Yichun Shi, Guoxian Song, Ümit Y. Ogras, Linjie Luo:
PanoHead: Geometry-Aware 3D Full-Head Synthesis in 360°. 20950-20959
- Hao Li, Xianxu Hou, Zepeng Huang, Linlin Shen:
StyleGene: Crossover and Mutation of Region-level Facial Genes for Kinship Face Synthesis. 20960-20969
- Mausoom Sarkar, Nikitha S. R., Mayur Hemani, Rishabh Jain, Balaji Krishnamurthy:
Parameter Efficient Local Implicit Image Function Network for Face Segmentation. 20970-20980
- Chang Yu, Xiangyu Zhu, Xiaomei Zhang, Zhaoxiang Zhang, Zhen Lei:
Graphics Capsule: Learning Hierarchical 3D Face Representations from 2D Images. 20981-20990
- Jingxiang Sun, Xuan Wang, Lizhen Wang , Xiaoyu Li, Yong Zhang, Hongwen Zhang, Yebin Liu:
Next3D: Generative Neural Texture Rasterization for 3D-Aware Head Avatars. 20991-21002
- Simon Giebenhain, Tobias Kirschstein, Markos Georgopoulos, Martin Rünz, Lourdes Agapito, Matthias Nießner:
Learning Neural Parametric Head Models. 21003-21012
- Rui Zhao, Wei Li, Zhipeng Hu, Lincheng Li, Zhengxia Zou, Zhenwei Shi, Changjie Fan:
Zero-Shot Text-to-Parameter Translation for Game Character Auto-Creation. 21013-21023
- Hsuan-I Ho , Lixin Xue, Jie Song, Otmar Hilliges:
Learning Locally Editable Virtual Humans. 21024-21035
- Yonggan Fu, Yuecheng Li, Chenghui Li, Jason M. Saragih, Peizhao Zhang, Xiaoliang Dai, Yingyan Celine Lin:
Auto-CARD: Efficient and Robust Codec Avatar Driving for Real-time Mobile Telepresence. 21036-21045
- Rotem Shalev-Arkushin, Amit Moryossef, Ohad Fried:
Ham2Pose: Animating Sign Language Notation into Pose Sequences. 21046-21056
- Yufeng Zheng, Wang Yifan, Gordon Wetzstein, Michael J. Black, Otmar Hilliges:
PointAvatar: Deformable Point-Based Head Avatars from Videos. 21057-21067
- Shuhong Chen, Kevin Zhang , Yichun Shi, Heng Wang, Yiheng Zhu, Guoxian Song, Sizhe An, Janus Kristjansson, Xiao Yang, Matthias Zwicker:
PAniC-3D: Stylized Single-view 3D Reconstruction from Portraits of Anime Characters. 21068-21077
- Zhiyang Guo , Wengang Zhou, Min Wang, Li Li, Houqiang Li:
HandNeRF: Neural Radiance Fields for Animatable Interacting Hands. 21078-21087
- Rishabh Jain, Krishna Kumar Singh, Mayur Hemani, Jingwan Lu, Mausoom Sarkar, Duygu Ceylan, Balaji Krishnamurthy:
VGFlow: Visibility guided Flow Network for Human Reposing. 21088-21097
- Kangkan Wang, Guofeng Zhang, Suxu Cong, Jian Yang:
Clothed Human Performance Capture with a Double-layer Neural Radiance Fields. 21098-21107
- Lixin Yang, Jian Xu, Licheng Zhong, Xinyu Zhan, Zhicheng Wang, Kejian Wu, Cewu Lu:
POEM: Reconstructing Hand in a Point Embedded Multi-view Stereo. 21108-21112
- Vinoj Jayasundara, Amit Agrawal, Nicolas Heron, Abhinav Shrivastava, Larry S. Davis:
FlexNeRF: Photorealistic Free-viewpoint Rendering of Moving Humans from Sparse Views. 21118-21127
- Chaoyang Wang, Lachlan Ewen MacDonald, László A. Jeni, Simon Lucey :
Flow Supervision for Deformable NeRF. 21128-21137
- Shaowei Liu, Saurabh Gupta, Shenlong Wang:
Building Rearticulable Models for Arbitrary 3D Objects from 4D Point Clouds. 21138-21147
- Hanbyel Cho, Yooshin Cho, Jaesung Ahn, Junmo Kim:
Implicit 3D Human Mesh Recovery using Consistency with Pose and Shape from Unseen-view. 21148-21158
- Jing Lin, Ailing Zeng, Haoqian Wang, Lei Zhang, Yu Li :
One-Stage 3D Whole-Body Mesh Recovery with Component Aware Transformer. 21159-21168
- Jihyun Lee, Minhyuk Sung , Honggyu Choi, Tae-Kyun Kim :
Im2Hands: Learning Attentive Implicit Representation of Interacting Two-Hand Shapes. 21169-21178
- Purva Tendulkar, Dídac Surís, Carl Vondrick:
FLEX: Full-Body Grasping Without Full-Body Grasps. 21179-21189
- Chen Bao, Helin Xu, Yuzhe Qin , Xiaolong Wang:
DexArt: Benchmarking Generalizable Dexterous Manipulation with Articulated Objects. 21190-21200
- Nick Heppert , Muhammad Zubair Irshad, Sergey Zakharov, Katherine Liu, Rares Andrei Ambrus, Jeannette Bohg , Abhinav Valada, Thomas Kollar:
CARTO: Category and Joint Agnostic Reconstruction of ARTiculated Objects. 21201-21210
- João Pedro Araújo, Jiaman Li, Karthik Vetrivel, Rishi Agarwal, Jiajun Wu, Deepak Gopinath, Alexander Clegg, C. Karen Liu :
CIRCLE: Capture In Rich Contextual Environments. 21211-21221



- Vickie Ye, Georgios Pavlakos, Jitendra Malik, Angjoo Kanazawa:
Decoupling Human and Camera Motion from Videos in the Wild. 21222-21232
- Han Xue, Wenqiang Xu, Jieyi Zhang, Tutan Tang, Yutong Li, Wenxin Du, Ruolin Ye, Cewu Lu:
GarmentTracking: Category-Level Garment Pose Tracking. 21233-21242
- Yilin Wen , Hao Pan, Lei Yang, Jia Pan, Taku Komura, Wenping Wang:
Hierarchical Temporal Transformer for 3D Hand Pose Estimation and Action Recognition from Egocentric RGB Videos. 21243-21253
- Zhongwei Qiu, Qiansheng Yang, Jian Wang, Haocheng Feng, Junyu Han, Errui Ding, Chang Xu, Dongmei Fu, Jingdong Wang:
PSVT: End-to-End Multi-Person 3D Pose and Shape Estimation with Progressive Video Transformers. 21254-21263
- Yulin Liu, Haoran Liu, Yingda Yin, Yang Wang, Baoquan Chen, He Wang:
Delving into Discrete Normalizing Flows on SO(3) Manifold for Probabilistic Rotation Modeling. 21264-21273
- Hemal Naik, Alex Hoi Hang Chan, Junran Yang, Mathilde Delacoux, Iain D. Couzin, Fumihiro Kano, Nagy Máté:
3D-POP - An Automated Annotation Approach to Facilitate Markerless 2D-3D Tracking of Freely Moving Birds with Marker-Based Motion Capture. 21274-21284
- Taeyeop Lee, Jonathan Tremblay, Valts Blukis, Bowen Wen, Byeong-Uk Lee, Inkyu Shin, Stan Birchfield, In So Kweon, Kuk-Jin Yoon:
TTA-COPE: Test-Time Adaptation for Category-Level Object Pose Estimation. 21285-21295
- Jingpei Lu, Florian Richter, Michael C. Yip:
Markerless Camera-to-Robot Pose Estimation via Self-Supervised Sim-to-Real Transfer. 21296-21306
- Tao Tan, Qiulei Dong:
SMOC-Net: Leveraging Camera Pose for Self-Supervised Monocular Object Pose Estimation. 21307-21316
- Fei Xue, Ignas Budvytis, Roberto Cipolla:
IMP: Iterative Matching and Pose Estimation with Adaptive Pooling. 21317-21326
- Benjamin T. Jones, Michael Hu, Milin Kodnongbua, Vladimir G. Kim, Adriana Schulz:
Self-Supervised Representation Learning for CAD. 21327-21336
- Xingzhe He, Gaurav Bharaj, David Ferman, Helge Rhodin, Pablo Garrido:
Few-Shot Geometry-Aware Keypoint Localization. 21337-21348
- Samarth Sinha, Jason Y. Zhang, Andrea Tagliasacchi, Igor Gilitschenski, David B. Lindell :
SparsePose: Sparse-View Camera Pose Regression and Refinement. 21349-21359
- Daniel Barath, Dmytro Mishkin, Michal Polic, Wolfgang Förstner, Jiri Matas:
A Large-Scale Homography Benchmark. 21360-21370
- Pattaramanee Arsomngern, Sarana Nutanong, Supasorn Suwajanakorn:
Learning Geometric-Aware Properties in 2D Representation Using Lightweight CAD Models, or Zero Real 3D Pairs. 21371-21381
- Yuang Wang, Xingyi He, Sida Peng, Haotong Lin, Hujun Bao, Xiaowei Zhou:
AutoRecon: Automated 3D Object Discovery and Reconstruction. 21382-21391
- Oleg Voynov, Gleb Bobrovskikh, Pavel A. Karpyshев, Saveliy Galochkin, Andrei-Timotei Ardelean, Arseniy Bozhenko, Ekaterina Karmanova, Pavel Kopanev, Yaroslav Labutin-Rymsho, Ruslan Rakhimov, Aleksandr Safin, Valerii Serpiva, Alexey Artemov, Evgeny Burnaev, Dzmitry Tsetserukou, Denis Zorin:
Multi-Sensor Large-Scale Dataset for Multi-View 3D Reconstruction. 21392-21403
- Zhixiang Min, Bingbing Zhuang, Samuel Schulter, Buyu Liu, Enrique Dunn, Manmohan Chandraker:
NeurOCS: Neural NOCS Supervision for Monocular 3D Object Localization. 21404-21414
- Botao Ye, Sifei Liu, Xuetong Li, Ming-Hsuan Yang:
Self-Supervised Super-Plane for Neural 3D Reconstruction. 21415-21424
- Ruoyu Wang, Zehao Yu, Shenghua Gao:
PlaneDepth: Self-Supervised Depth Estimation via Orthogonal Planes. 21425-21434
- Byeong-Uk Lee, Jianming Zhang, Yannick Hold-Geoffroy, In So Kweon:
Single View Scene Scale Estimation using Scale Field. 21435-21444
- Shaohui Liu, Yifan Yu, Rémi Pautrat, Marc Pollefeys, Viktor Larsson:
3D Line Mapping Revisited. 21445-21455
- Xin Huang, Qi Zhang, Ying Feng, Hongdong Li, Qing Wang:
Inverting the Imaging Process by Learning an Implicit Camera Model. 21456-21465
- Sergio Izquierdo, Javier Civera:
SfM-TTR: Using Structure from Motion for Test-Time Refinement of Single-View Depth Networks. 21466-21476
- Luigi Piccinelli, Christos Sakaridis , Fisher Yu:
iDisc: Internal Discretization for Monocular Depth Estimation. 21477-21487
- Hadi Alzayer, Abdullah Abuolaim, Leung Chun Chan, Yang Yang, Ying Chen Lou, Jia-Bin Huang, Abhishek Kar:
DC²: Dual-Camera Defocus Control by Learning to Refocus. 21488-21497
- Jialiang Wang, Daniel Scharstein, Akash Bapat, Kevin Blackburn-Matzen, Matthew Yu, Jonathan Lehman, Suhib Alsisan, Yanghan Wang, Sam S. Tsai, Jan-Michael Frahm, Zijian He, Peter Vajda, Michael F. Cohen, Matt Uyttendaele:
A Practical Stereo Depth System for Smart Glasses. 21498-21507



- Zhe Zhang, Rui Peng, Yuxi Hu, Ronggang Wang:
GeoMVSNet: Learning Multi-View Stereo with Geometry Perception. 21508-21518
- Yichen Guo, Mai Xu, Lai Jiang, Leonid Sigal, Yunjin Chen:
DINN360: Deformable Invertible Neural Network for Latitude-aware 360° Image Rescaling. 21519-21528
- Sheng Xie, Daochuan Wang, Yunhui Liu:
OmniVidar: Omnidirectional Depth Estimation from Multi-Fisheye Images. 21529-21538
- Rui Li, Dong Gong, Wei Yin, Hao Chen, Yu Zhu, Kaixuan Wang, Xiaozhi Chen , Jinqiu Sun, Yanning Zhang:
Learning to Fuse Monocular and Multi-view Cues for Multi-frame Depth Estimation in Dynamic Scenes. 21539-21548
- Marius Memmel, Roman Bachmann, Amir Zamir:
Modality-invariant Visual Odometry for Embodied Vision. 21549-21559
- Ziqin Wang, Bowen Cheng , Lichen Zhao, Dong Xu, Yang Tang, Lu Sheng:
VL-SAT: Visual-Linguistic Semantics Assisted Training for 3D Semantic Scene Graph Prediction in Point Cloud. 21560-21569
- Kaixin Xiong, Shi Gong, Xiaoqing Ye, Xiao Tan, Ji Wan, Errui Ding, Jingdong Wang, Xiang Bai:
CAPE: Camera View Position Embedding for Multi-View 3D Object Detection. 21570-21579
- Chengjian Feng, Zequn Jie, Yujie Zhong, Xiangxiang Chu, Lin Ma:
AeDet: Azimuth-Invariant Multi-View 3D Object Detection. 21580-21588
- Zekun Zhang, Minh Hoai:
Object Detection with Self-Supervised Scene Adaptation. 21589-21599
- Zijian Zhu, Yichi Zhang, Hai Chen, Yinpeng Dong, Shu Zhao, Wenbo Ding, Jiachen Zhong, Shibao Zheng:
Understanding the Robustness of 3D Object Detection with Bird'View Representations in Autonomous Driving. 21600-21610
- Ruihao Wang, Jian Qin, Kaiying Li, Yaochen Li, Dong Cao, Jintao Xu:
BEV-LaneDet: An Efficient 3D Lane Detection Based on Virtual Camera via Key-Points. 21611-21620
- Lei Yang , Kaicheng Yu, Tao Tang, Jun Li, Kun Yuan, Li Wang, Xinyu Zhang, Peng Chen:
BEVHeight: A Robust Framework for Vision-based Roadside 3D Object Detection. 21611-21620
- Florian Fervers, Sebastian Bullinger, Christoph Bodensteiner, Michael Arens, Rainer Stiefelhagen:
Uncertainty-Aware Vision-Based Metric Cross-View Geolocalization. 21621-21631
- Paul-Edouard Sarlin, Daniel DeTone, Tsun-Yi Yang, Armen Avetisyan, Julian Straub, Tomasz Malisiewicz, Samuel Rota Bulò, Richard A. Newcombe, Peter Kontschieder, Vasileios Balntas:
OrienterNet: Visual Localization in 2D Public Maps with Neural Matching. 21632-21642
- Yang Jiao, Zequn Jie, Shaoxiang Chen, Jingjing Chen , Lin Ma, Yu-Gang Jiang:
MSMDFusion: Fusing LiDAR and Camera at Multiple Scales with Multi-Depth Seeds for 3D Object Detection. 21643-21652
- Hai Wu, Chenglu Wen, Shaoshuai Shi, Xin Li, Cheng Wang:
Virtual Sparse Convolution for Multimodal 3D Object Detection. 21653-21662
- Wei Lin , Antoni B. Chan:
Optimal Transport Minimization: Crowd Localization on Density Maps for Semi-Supervised Counting. 21663-21673
- Yukang Chen, Jianhui Liu, Xiangyu Zhang, Xiaojuan Qi, Jiaya Jia:
VoxelNeXt: Fully Sparse VoxelNet for 3D Object Detection and Tracking. 21674-21683
- Oren Shtrut, Yizhak Ben-Shabat, Ayellet Tal:
GraVoS: Voxel Selection for 3D Point-Cloud Detection. 21684-21693
- Jiale Li, Hang Dai, Hao Han, Yong Ding:
MSeg3D: Multi-Modal 3D Semantic Segmentation for Autonomous Driving. 21694-21704
- Lingdong Kong , Jiawei Ren, Liang Pan, Ziwei Liu:
LaserMix for Semi-Supervised LiDAR Semantic Segmentation. 21706-21716
- Zaiwei Zhang, Min Bai , Li Erran Li:
Implicit Surface Contrastive Clustering for LiDAR Point Clouds. 21716-21725
- Gengxin Liu, Qian Sun, Haibin Huang, Chongyang Ma, Yulan Guo, Li Yi, Hui Huang, Ruizhen Hu:
Semi-Weakly Supervised Object Kinematic Motion Prediction. 21726-21735
- Minghua Liu, Yinhao Zhu, Hong Cai, Shizhong Han, Zhan Ling, Fatih Porikli, Hao Su:
PartSLIP: Low-Shot Part Segmentation for 3D Point Clouds via Pretrained Image-Language Models. 21736-21746
- Yurui Zhu, Tianyu Wang, Xueyang Fu, Xuanyu Yang, Xin Guo, Jifeng Dai, Yu Qiao, Xiaowei Hu:
Learning Weather-General and Weather-Specific Features for Image Restoration Under Multiple Adverse Weather Conditions. 21747-21758
- Yuwei Yang, Munawar Hayat, Zhao Jin, Chao Ren, Yinjie Lei:
Geometry and Uncertainty-Aware 3D Point Cloud Class-Incremental Semantic Segmentation. 21759-21768
- Renrui Zhang, Liuhui Wang, Yu Qiao, Peng Gao, Hongsheng Li :
Learning 3D Representations from 2D Pre-Trained Models via Image-to-Point Masked Autoencoders. 21769-21780



- Xinglin Li, Jiajing Chen, Jinhui Ouyang, Hanhui Deng , Senem Velipasalar, Di Wu:
ToThePoint: Efficient Contrastive Learning of 3D Point Clouds via Recycling. 21781-21790
- Linfeng Zhang, Runpei Dong , Hung-Shuo Tai, Kaisheng Ma:
PointDistiller: Structured Knowledge Distillation Towards Efficient and Compact 3D Detection. 21791-21801
- Wenxuan Wu, Fuxin Li, Qi Shan:
PointConvFormer: Revenge of the Point-based Convolution. 21802-21813
- Jinyoung Park, Sanghyeok Lee, Sihyeon Kim, Yunyang Xiong, Hyunwoo J. Kim:
Self-Positioning Point-Based Transformer for Point Cloud Understanding. 21814-21823
- Fuchen Long, Ting Yao, Zhaofan Qiu, Lusong Li, Tao Mei:
PointClustering: Unsupervised Point Cloud Pre-training using Transformation Invariance in Clustering. 21824-21834
- Puhua Jiang, Mingze Sun, Ruqi Huang:
Neural Intrinsic Embedding for Non-Rigid Point Cloud Matching. 21835-21845
- Ting Yao, Yehao Li, Yingwei Pan, Tao Mei:
HGNet: Learning Hierarchical Geometry from Points, Edges, and Surfaces. 21846-21855
- Meng Wang, Yu-Shen Liu, Yue Gao, Kanle Shi, Yi Fang, Zhizhong Han:
LP-DIF: Learning Local Pattern-Specific Deep Implicit Function for 3D Objects and Scenes. 21856-21865
- Paul Roetzer, Zorah Lähner , Florian Bernard:
Conjugate Product Graphs for Globally Optimal 2D-3D Shape Matching. 21866-21875
- Sisi You, Hantao Yao, Bing-Kun Bao, Changsheng Xu:
UTM: A Unified Multiple Object Tracking Model with Identity-Aware Feature Enhancement. 21876-21886
- Jongmin Lee, Byungjin Kim, Seungwook Kim, Minsu Cho:
Learning Rotation-Equivariant Features for Visual Correspondence. 21887-21897
- Jiahuan Yu, Jiahao Chang, Jianfeng He, Tianzhu Zhang, Jiyang Yu, Feng Wu:
Adaptive Spot-Guided Transformer for Consistent Local Feature Matching. 21898-21908
- Shengjie Zhu , Xiaoming Liu:
PMatch: Paired Masked Image Modeling for Dense Geometric Matching. 21909-21918
- Gangwei Xu, Xianqi Wang, Xiaohuan Ding, Xin Yang:
Iterative Geometry Encoding Volume for Stereo Matching. 21919-21928
- Chitturi Sidhartha, Lalit Manam, Venu Madhav Govindu:
Adaptive Annealing for Robust Geometric Estimation. 21929-21939
- Jun Nagata, Yusuke Sekikawa:
Tangentially Elongated Gaussian Belief Propagation for Event-Based Incremental Optical Flow Estimation. 21940-21949
- Yifan Lu, Jiayi Ma, Leyuan Fang, Xin Tian, Junjun Jiang:
Robust and Scalable Gaussian Process Regression and Its Applications. 21950-21959
- Yunze Man, Liang-Yan Gui, Yu-Xiong Wang:
BEV-Guided Multi-Modality Fusion for Driving Perception. 21960-21969
- Shixiang Tang, Cheng Chen, Qingsong Xie, Meilin Chen, Yizhou Wang, Yuanzheng Ci, Lei Bai, Feng Zhu, Haiyang Yang, Li Yi, Rui Zhao, Wanli Ouyang:
HumanBench: Towards General Human-Centric Perception with Projector Assisted Pretraining. 21970-21982
- Xiaosong Jia, Penghao Wu, Li Chen, Jiangwei Xie, Conghui He, Junchi Yan, Hongyang Li:
Think Twice before Driving: Towards Scalable Decoders for End-to-End Autonomous Driving. 21983-21994
- Xishun Wang, Tong Su, Fang Da, Xiaodong Yang:
ProphNet: Efficient Agent-Centric Motion Forecasting with Anchor-Informed Proposals. 21995-22003
- Sean Kulinski, Nicholas R. Waytowich, James Z. Hare, David I. Inouye:
StarCraftImage: A Dataset For Prototyping Spatial Reasoning Methods For Multi-Agent Environments. 22004-22013
- Jianhua Sun, Yuxuan Li, Liang Chai, Cewu Lu:
Stimulus Verification is a Universal and Effective Sampler in Multi-modal Human Trajectory Prediction. 22014-22023
- Chen Wang, Dasong Gao, Kuan Xu, Junyi Geng, Yaoyu Hu, Yuheng Qiu, Bowen Li, Fan Yang, Brady G. Moon, Abhinav Pandey, Aryan, Jiahe Xu, Tianhao Wu, Haonan He, Daning Huang, Zhongqiang Ren, Shibo Zhao, Taimeng Fu, Pranay Reddy, Xiao Lin, Wenshan Wang, Jingnan Shi, Rajat Talak, Kun Cao, Yi Du, Han Wang, Huai Yu, Shanzhao Wang, Siyu Chen, Ananth Kashyap, Rohan Bandaru, Karthik Dantu, Jiajun Wu, Lihua Xie, Luca Carbone, Marco Hutter, Sebastian A. Scherer :
PyPose: A Library for Robot Learning with Physics-based Optimization. 22024-22034
- Xin Cai, Jiabei Zeng, Shiguang Shan, Xilin Chen:
Source-Free Adaptive Gaze Estimation by Uncertainty Reduction. 22035-22045
- Chunming He, Kai Li, Yachao Zhang, Longxiang Tang, Yulun Zhang, Zhenhua Guo, Xiu Li:
Camouflaged Object Detection with Feature Decomposition and Edge Reconstruction. 22046-22055



- Yuang Zhang, Tiancai Wang, Xiangyu Zhang:
MOTRv2: Bootstrapping End-to-End Multi-Object Tracking by Pretrained Object Detectors. 22056-22065
- Kang Ma, Ying Fu, Dezhi Zheng, Chunshui Cao, Xuecai Hu, Yongzhen Huang:
Dynamic Aggregated Network for Gait Recognition. 22076-22085
- Zhijun Zhai, Jianhui Zhao, Chengjiang Long, Wenju Xu, Shuangjiang He , Huijuan Zhao:
Feature Representation Learning with Adaptive Displacement Generation and Transformer Fusion for Micro-Expression Recognition. 22086-22095
- Bowen Zhang, Chenyang Qi, Pan Zhang, Bo Zhang, HsiangTao Wu, Dong Chen, Qifeng Chen, Yong Wang, Fang Wen:
MetaPortrait: Identity-Preserving Talking Head Generation with Fast Personalized Adaptation. 22096-22105
- Yansong Tang, Jinpeng Liu, Aoyang Liu, Bin Yang, Wenxun Dai, Yongming Rao, Jiwen Lu, Jie Zhou, Xiu Li:
FLAG3D: A 3D Fitness Activity Dataset with Language Instruction. 22106-22117
- Haocong Rao , Chunyan Miao:
TranSG: Transformer-Based Skeleton Graph Prototype Contrastive Learning with Structure-Trajectory Prompted Reconstruction for Person Re-Identification. 22118-22128
- Kuan-Chieh Wang, Zhenzhen Weng , Maria Xenochristou, João Pedro Araújo, Jeffrey Gu, C. Karen Liu , Serena Yeung :
NeMo: 3D Neural Motion Fields from Multiple Video Instances of the Same Action. 22129-22138
- Etienne Meunier, Patrick Bouthemy:
Unsupervised Space-Time Network for Temporally-Consistent Segmentation of Multiple Motions. 22139-22148
- Haiyang Mei, Zuowen Wang, Xin Yang, Xiaopeng Wei, Tobi Delbrück:
Deep Polarization Reconstruction with PDAVIS Events. 22149-22158
- Zhiyang Yu , Yu Zhang, Dongqing Zou, Xijun Chen, Jimmy S. Ren, Shunqing Ren:
Range-nullspace Video Frame Interpolation with Focalized Motion Estimation. 22159-22168
- Kun Zhou , Wenbo Li, Xiaoguang Han, Jiangbo Lu:
Exploring Motion Ambiguity and Alignment for High-Quality Video Frame Interpolation. 22169-22179
- Yakun Chang, Chu Zhou, Yuchen Hong, Liwen Hu, Chao Xu, Tiejun Huang, Boxin Shi:
1000 FPS HDR Video with a Spike-RGB Hybrid Camera. 22180-22190
- Jinshan Pan, Boming Xu, Jiangxin Dong, Jianjun Ge, Jinhui Tang:
Deep Discriminative Spatial and Temporal Network for Efficient Video Deblurring. 22191-22200
- Nancy Mehta , Akshay Dudhane, Subrahmanyam Murala, Syed Waqas Zamir, Salman H. Khan, Fahad Shahbaz Khan:
Gated Multi-Resolution Transfer Network for Burst Restoration and Enhancement. 22201-22210
- Qingsen Yan, Weiye Chen, Song Zhang, Yu Zhu, Jinqiu Sun, Yanning Zhang:
A Unified HDR Imaging Method with Pixel and Patch Level. 22211-22220
- Nikolai Kalischek, Rodrigo Caye Daudt, Torben Peters, Reinhard Furrer, Jan D. Wegner, Konrad Schindler:
BiasBed - Rigorous Texture Bias Evaluation. 22221-22230
- Cheng Guo , Leidong Fan, Ziyu Xue, Xiuhua Jiang:
Learning a Practical SDR-to-HDR TV Up-conversion using New Dataset and Degradation Models. 22231-22241
- Haoyu Chen , Zhihua Wang, Yang Yang, Qilin Sun, Kede Ma :
Learning a Deep Color Difference Metric for Photographic Images. 22242-22251
- Zhenqi Fu, Yan Yang, Xiaotong Tu, Yue Huang, Xinghao Ding, Kai-Kuang Ma:
Learning a Simple Low-Light Image Enhancer from Paired Low-Light Instances. 22252-22261
- Yubo Dong, Dahua Gao, Tian Qiu, Yuyan Li , Minxi Yang, Guangming Shi:
Residual Degradation Learning Unfolding Framework with Mixing Priors Across Spectral and Spatial for Compressive Spectral Imaging. 22262-22271
- Wen-jin Guo, Weiyi Xie, Kai Jiang, Yunsong Li, Jie Lei, Leyuan Fang:
Toward Stable, Interpretable, and Lightweight Hyperspectral Super-Resolution. 22272-22281
- Ruiqi Wu, Zheng-Peng Duan, Chun-Le Guo, Zhi Chai, Chongyi Li:
RIDCP: Revitalizing Real Image Dehazing via High-Quality Codebook Priors. 22282-22291
- Yohan Poirier-Ginter, Jean-François Lalonde:
Robust Unsupervised StyleGAN Image Restoration. 22292-22301
- Kai Zhao, Kun Yuan, Ming Sun, Mading Li, Xing Wen:
Quality-aware Pretrained Models for Blind Image Quality Assessment. 22302-22313
- Haina Qin, Longfei Han, Weihua Xiong, Juan Wang, Wentao Ma, Bing Li, Weiming Hu:
Learning to Exploit the Sequence-Specific Prior Knowledge for Image Processing Pipelines Optimization. 22314-22323
- Eirikur Agustsson, David Minnen, George Toderici, Fabian Mentzer:
Multi-Realism Image Compression with a Conditional Generator. 22324-22333
- Jeongsoo Park, Justin Johnson:
RGB No More: Minimally-Decoded JPEG Vision Transformers. 22334-22346
- Michael Bernasconi, Abdelaziz Djelouah, Farnood Salehi, Markus Gross , Christopher Schroers:
Kernel Aware Resampler. 22347-22355



- Chenyang Wang, Junjun Jiang, Zhiwei Zhong, Xianming Liu:
Spatial-Frequency Mutual Learning for Face Super-Resolution. 22356-22366
- Xiangyu Chen, Xintao Wang, Jiantao Zhou, Yu Qiao, Chao Dong:
Activating More Pixels in Image Super-Resolution Transformer. 22367-22377
- Hang Wang, Xuanhong Chen, Bingbing Ni, Yutian Liu, Jinfan Liu:
Omni Aggregation Networks for Lightweight Image Super-Resolution. 22378-22387
- Ran Yi, Haoyuan Tian, Zhihao Gu, Yu-Kun Lai, Paul L. Rosin:
Towards Artistic Image Aesthetics Assessment: a Large-scale Dataset and a New Method. 22388-22397
- Luwen Duan, Min Wu, Lijian Mao, Jun Yin, Jianping Xiong, Xi Li:
RWSC-Fusion: Region-Wise Style-Controlled Fusion Network for the Prohibited X-ray Security Image Synthesis. 22398-22407
- Thuan Hoang Nguyen , Thanh Van Le, Anh Tran:
Efficient Scale-Invariant Generator with Column-Row Entangled Pixel Synthesis. 22408-22417
- Chang Jiang, Fei Gao, Biao Ma, Yuhao Lin, Nannan Wang, Gang Xu:
Masked and Adaptive Transformer for Exemplar Based Image Translation. 22418-22427
- Shaoan Xie, Zhifei Zhang, Zhe Lin, Tobias Hinz, Kun Zhang:
SmartBrush: Text and Shape Guided Object Inpainting with Diffusion Model. 22428-22437
- Bin Fu, Junjun He, Jianjun Wang, Yu Qiao:
Neural Transformation Fields for Arbitrary-Styled Font Generation. 22438-22447
- Jizhizi Li, Jing Zhang, Dacheng Tao:
Referring Image Matting. 22448-22457
- Vittorio Pippi, Silvia Cascianelli, Rita Cucchiara:
Handwritten Text Generation from Visual Archetypes. 22458-22467
- Yu Zeng, Zhe Lin, Jianming Zhang, Qing Liu, John P. Collomosse, Jason Kuen, Vishal M. Patel:
SceneComposer: Any-Level Semantic Image Synthesis. 22468-22478
- Yufei Ye, Xuetong Li, Abhinav Gupta, Shalini De Mello, Stan Birchfield, Jiaming Song, Shubham Tulsiani, Sifei Liu:
Affordance Diffusion: Synthesizing Hand-Object Interactions. 22479-22489
- Guangcong Zheng, Xianpan Zhou, Xuewei Li , Zhongang Qi, Ying Shan, Xi Li:
LayoutDiffusion: Controllable Diffusion Model for Layout-to-Image Generation. 22490-22499
- Nataniel Ruiz, Yuanzhen Li, Varun Jampani, Yael Pritch, Michael Rubinstein, Kfir Aberman:
DreamBooth: Fine Tuning Text-to-Image Diffusion Models for Subject-Driven Generation. 22500-22510
- Yuheng Li, Haotian Liu, Qingyang Wu, Fangzhou Mu, Jianwei Yang, Jianfeng Gao, Chunyuan Li, Yong Jae Lee:
GLIGEN: Open-Set Grounded Text-to-Image Generation. 22511-22521
- Patrick Schramowski, Manuel Brack, Björn Deisereth, Kristian Kersting:
Safe Latent Diffusion: Mitigating Inappropriate Degeneration in Diffusion Models. 22522-22531
- Bram Wallace, Akash Gokul, Nikhil Naik:
EDICT: Exact Diffusion Inversion via Coupled Transformations. 22532-22541
- Hyungjin Chung, Dohoon Ryu, Michael T. McCann, Marc Louis Klasky, Jong Chul Ye:
Solving 3D Inverse Problems Using Pre-Trained 2D Diffusion Models. 22542-22551
- Xingyi Yang , Daquan Zhou , Jiashi Feng, Xinchao Wang:
Diffusion Probabilistic Model Made Slim. 22552-22562
- Andreas Blattmann, Robin Rombach, Huan Ling, Tim Dockhorn, Seung Wook Kim, Sanja Fidler, Karsten Kreis:
Align Your Latents: High-Resolution Video Synthesis with Latent Diffusion Models. 22563-22575
- Ze Wang, Jiang Wang, Zicheng Liu, Qiang Qiu:
Binary Latent Diffusion. 22576-22585
- Zhiliang Wu , Hanyu Xuan, Changchang Sun, Weili Guan, Kang Zhang, Yan Yan:
Semi-Supervised Video Inpainting with Cycle Consistency Constraints. 22586-22595
- Mengqi Huang, Zhendong Mao, Zhuowei Chen, Yongdong Zhang:
Towards Accurate Image Coding: Improved Autoregressive Image Generation with Dynamic Vector Quantization. 22596-22605
- Chong Mou, Youmin Xu, Jiechong Song, Chen Zhao, Bernard Ghanem, Jian Zhang:
Large-Capacity and Flexible Video Steganography via Invertible Neural Network. 22606-22615
- Jiahao Li, Bin Li, Yan Lu:
Neural Video Compression with Diverse Contexts. 22616-22626
- Yubin Hu , Yuze He, Yanghao Li, Jisheng Li, Yuxing Han, Jiangtao Wen, Yong-Jin Liu:
Efficient Semantic Segmentation by Altering Resolutions for Compressed Videos. 22627-22637
- Bin Xia, Jingwen He, Yulun Zhang, Yitong Wang, Yapeng Tian, Wenming Yang, Luc Van Gool:
Structured Sparsity Learning for Efficient Video Super-Resolution. 22638-22647
- Yihao Chen, Xianbiao Qi, Jianan Wang, Lei Zhang:
DisCo-CLIP: A Distributed Contrastive Loss for Memory Efficient CLIP Training. 22648-22657



- Chong Yu, Tao Chen, Zhongxue Gan, Jiayuan Fan:
Boost Vision Transformer with GPU-Friendly Sparsity and Quantization. 22658-22668
- Fan Bao, Shen Nie, Kaiwen Xue, Yue Cao, Chongxuan Li, Hang Su, Jun Zhu:
All are Worth Words: A ViT Backbone for Diffusion Models. 22669-22679
- Cong Wei, Brendan Duke, Ruowei Jiang, Parham Aarabi, Graham W. Taylor, Florian Shkurti:
Sparsifiner: Learning Sparse Instance-Dependent Attention for Efficient Vision Transformers. 22680-22689
- Bonan Li, Yinhan Hu, Xuecheng Nie, Congying Han, Xiangjian Jiang, Tiande Guo, Luoqi Liu:
DropKey for Vision Transformer. 22700-22709
- Zijiao Chen, Jiaxin Qing, Tiange Xiang, Wan Lin Yue, Juan Helen Zhou:
Seeing Beyond the Brain: Conditional Diffusion Model with Sparse Masked Modeling for Vision Decoding. 22710-22720
- Rui Tian, Zuxuan Wu, Qi Dai, Han Hu, Yu Qiao, Yu-Gang Jiang:
ResFormer: Scaling ViTs with Multi-Resolution Training. 22721-22731
- Hongwei Xue, Peng Gao, Hongyang Li, Yu Qiao, Hao Sun, Houqiang Li, Jiebo Luo :
Stare at What You See: Masked Image Modeling without Reconstruction. 22732-22741
- Kai Chen, Zhili Liu, Lanqing Hong, Hang Xu, Zhenguo Li, Dit-Yan Yeung:
Mixed Autoencoder for Self-Supervised Visual Representation Learning. 22742-22751
- Jiawei Feng, Ancong Wu, Wei-Shi Zheng:
Shape-Erased Feature Learning for Visible-Infrared Person Re-Identification. 22752-22761
- Marvin Eisenberger, Aysim Toker, Laura Leal-Taixé, Daniel Cremers:
G-MSM: Unsupervised Multi-Shape Matching with Graph-Based Affinity Priors. 22762-22772
- Fei Du, Jianlong Yuan, Zhibin Wang, Fan Wang:
Efficient Mask Correction for Click-Based Interactive Image Segmentation. 22773-22782
- Chaofan Zheng, Xinyu Lyu, Lianli Gao, Bo Dai, Jingkuan Song:
Prototype-Based Embedding Network for Scene Graph Generation. 22783-22792
- Yue Qiu, Yanjun Sun, Fumiya Matsuzawa, Kenji Iwata, Hirokatsu Kataoka:
Graph Representation for Order-aware Visual Transformation. 22793-22802
- Sayak Nag, Kyle Min, Subarna Tripathi, Amit K. Roy-Chowdhury:
Unbiased Scene Graph Generation in Videos. 22803-22813
- Paul Micaelli, Arash Vahdat, Hongxu Yin, Jan Kautz, Pavlo Molchanov:
Recurrence without Recurrence: Stable Video Landmark Detection with Deep Equilibrium Models. 22814-22825
- Fei Xie, Lei Chu, Jiahao Li, Yan Lu, Chao Ma:
VideoTrack: Learning to Track Objects via Video Transformer. 22826-22835
- Pavel Tokmakov, Jie Li, Adrien Gaidon:
Breaking the "Object" in Video Object Segmentation. 22836-22845
- Shengyang Sun, Xiaojin Gong:
Hierarchical Semantic Contrast for Scene-aware Video Anomaly Detection. 22846-22856
- Lei Ke, Martin Danelljan, Henghui Ding , Yu-Wing Tai, Chi-Keung Tang, Fisher Yu:
Mask-Free Video Instance Segmentation. 22857-22866
- Ryuhei Hamaguchi , Yasutaka Furukawa, Masaki Onishi, Ken Sakurada:
Hierarchical Neural Memory Network for Low Latency Event Processing. 22867-22876
- Orcun Cetintas, Guillem Brasó, Laura Leal-Taixé:
Unifying Short and Long-Term Tracking with Graph Hierarchies. 22877-22887
- Jaehoon Yoo, Semin Kim, Doyup Lee, Chiheon Kim, Seunghoon Hong:
Towards End-to-End Generative Modeling of Long Videos with Memory-Efficient Bidirectional Transformers. 22888-22897
- Tsu-Jui Fu, Linjie Li, Zhe Gan, Kevin Lin, William Yang Wang, Lijuan Wang, Zicheng Liu:
An Empirical Study of End-to-End Video-Language Transformers with Masked Visual Modeling. 22898-22909
- Chao Huang, Yapeng Tian, Anurag Kumar, Chenliang Xu:
Egocentric Audio-Visual Object Localization. 22910-22921
- Paul Hongsuck Seo, Arsha Nagrani, Cordelia Schmid:
AVFormer: Injecting Vision into Frozen Speech Models for Zero-Shot AV-ASR. 22922-22931
- Junhua Liao , Haihan Duan, Kanghui Feng, Wanbing Zhao, Yanbing Yang, Liangyin Chen:
A Light Weight Model for Active Speaker Detection. 22932-22941
- Tiantian Geng, Teng Wang, Jinming Duan, Runmin Cong, Feng Zheng:
Dense-Localizing Audio-Visual Events in Untrimmed Videos: A Large-Scale Benchmark and Baseline. 22942-22951
- Wei Lin, Muhammad Jehanzeb Mirza, Mateusz Kozinski, Horst Possegger, Hilde Kuehne, Horst Bischof :
Video Test-Time Adaptation for Action Recognition. 22952-22961
- Ryo Hachiuma, Fumiaki Sato, Taiki Sekii:
Unified Keypoint-Based Action Recognition Framework via Structured Keypoint Pooling. 22962-22971



- Zhipeng Bao, Pavel Tokmakov, Yu-Xiong Wang, Adrien Gaidon, Martial Hebert:
Object Discovery from Motion-Guided Tokens. 22972-22981
- Chen Zhao, Dawei Du, Anthony Hoogs, Christopher Funk :
Open Set Action Recognition via Multi-Label Evidential Learning. 22982-22991
- Mamshad Nayeem Rizve, Gaurav Mittal, Ye Yu, Matthew Hall, Sandra Sajeev, Mubarak Shah, Mei Chen:
PivoTAL: Prior-Driven Supervision for Weakly-Supervised Temporal Action Localization. 22992-23002
- Jingqiu Zhou, Linjiang Huang, Liang Wang, Si Liu, Hongsheng Li :
Improving Weakly Supervised Temporal Action Localization by Bridging Train-Test Gap in Pseudo Labels. 23003-23012
- Wei Ji, Renjie Liang, Zhedong Zheng , Wenqiao Zhang, Shengyu Zhang, Juncheng Li, Mengze Li, Tat-Seng Chua:
Are Binary Annotations Sufficient? Video Moment Retrieval via Hierarchical Uncertainty-based Active Learning. 23013-23022
- WonJun Moon, Sangeek Hyun, Sanguk Park, Dongchan Park, Jae-Pil Heo:
Query - Dependent Video Representation for Moment Retrieval and Highlight Detection. 23023-23033
- Syed Talal Wasim , Muzammal Naseer, Salman H. Khan, Fahad Shahbaz Khan, Mubarak Shah:
Vita-CLIP: Video and text adaptive CLIP via Multimodal Prompting. 23034-23044
- Dezhao Luo, Jiabo Huang, Shaogang Gong, Hailin Jin, Yang Liu:
Towards Generalisable Video Moment Retrieval: Visual-Dynamic Injection to Image-Text Pre-Training. 23045-23055
- Abhay Zala, Jaemin Cho, Satwik Kottur, Xilun Chen, Barlas Oguz, Yashar Mehdad, Mohit Bansal:
Hierarchical Video-Moment Retrieval and Step-Captioning. 23056-23065
- Kumar Ashutosh, Rohit Girdhar, Lorenzo Torresani, Kristen Grauman:
HierVL: Learning Hierarchical Video-Language Embeddings. 23066-23078
- Ziyun Zeng, Yuying Ge, Xihui Liu, Bin Chen, Ping Luo, Shu-Tao Xia, Yixiao Ge:
Learning Transferable Spatiotemporal Representations from Natural Script Knowledge. 23079-23089
- Mengze Li, Han Wang, Wenqiao Zhang, Jiaxu Miao, Zhou Zhao, Shengyu Zhang, Wei Ji, Fei Wu:
WINNER: Weakly-supervised hierarchical decompositioN and alignMent for spatio-temporal video gRounding. 23090-23099
- Zihang Lin, Chaolei Tan, Jian-Fang Hu, Zhi Jin , Tiancai Ye, Wei-Shi Zheng:
Collaborative Static and Dynamic Vision-Language Streams for Spatio-Temporal Video Grounding. 23100-23109
- Davide Moltisanti , Frank Keller, Hakan Bilen , Laura Sevilla-Lara:
Learning Action Changes by Measuring Verb-Adverb Textual Relationships. 23110-23118
- Linjie Li, Zhe Gan, Kevin Lin, Chung-Ching Lin, Zicheng Liu, Ce Liu, Lijuan Wang:
LAVENDER: Unifying Video-Language Understanding as Masked Language Modeling. 23119-23129
- Lijin Yang, Quan Kong, Hsuan-Kung Yang, Wadim Kehl, Yoichi Sato, Norimasa Kobori:
DeCo: Decomposition and Reconstruction for Compositional Temporal Grounding via Coarse-to-Fine Contrastive Ranking. 23130-23140
- Jiangbin Zheng, Yile Wang, Cheng Tan, Siyuan Li, Ge Wang, Jun Xia, Yidong Chen, Stan Z. Li:
CVT-SLR: Contrastive Visual-Textual Transformation for Sign Language Recognition with Variational Alignment. 23141-23150
- Li Zhou, Zikun Zhou, Kaige Mao, Zhenyu He:
Joint Visual Grounding and Tracking with Natural Language Specification. 23151-23160
- Teng Wang, Yixiao Ge, Feng Zheng, Ran Cheng, Ying Shan, Xiaohu Qie, Ping Luo:
Accelerating Vision-Language Pretraining with Free Language Modeling. 23161-23170
- Samir Yitzhak Gadre, Mitchell Wortsman, Gabriel Ilharco, Ludwig Schmidt, Shuran Song :
CoWs on Pasture: Baselines and Benchmarks for Language-Driven Zero-Shot Object Navigation. 23171-23181
- Brandon Clark, Alec Kerrigan, Parth Parag Kulkarni, Vicente Vivanco Cepeda, Mubarak Shah:
Where We Are and What We're Looking At: Query Based Worldwide Image Geo-localization Using Hierarchies and Scenes. 23182-23190
- Zhou Yu, Lixiang Zheng, Zhou Zhao, Fei Wu, Jianping Fan, Kui Ren, Jun Yu:
ANetQA: A Large-scale Benchmark for Fine-grained Compositional Reasoning over Untrimmed Videos. 23191-23200
- Arjun R. Akula, Brendan Driscoll, Pradyumna Narayana, Soravit Changpinyo, Zhiwei Jia, Suyash Damle , Garima Pruthi, Sugato Basu, Leonidas J. Guibas, William T. Freeman, Yuanzhen Li, Varun Jampani:
MetaCLUE: Towards Comprehensive Visual Metaphors Research. 23201-23211
- Jingyang Huo, Qiang Sun, Boyan Jiang, Haitao Lin, Yanwei Fu :
GeoVNL: Learning Geometry-Enhanced Visual Representation with Slot Attention for Vision-and-Language Navigation. 23212-23221
- Junfan Lin, Jianlong Chang, Lingbo Liu, Guanbin Li, Liang Lin, Qi Tian, Chang Wen Chen:
Being Comes from Not-Being: Open-Vocabulary Text-to-Motion Generation with Wordless Training. 23222-23231
- Adrian Bulat, Georgios Tzimiropoulos:
LASP: Text-to-Text Optimization for Language-Aware Soft Prompting of Vision & Language Models. 23232-23241



- Jinpeng Wang, Pan Zhou , Mike Zheng Shou, Shuicheng Yan:
Position-Guided Text Prompt for Vision-Language Pre-Training. 23242-23251
- Qu Tang, Xiangyu Zhu, Zhen Lei, Zhaoxiang Zhang:
Intrinsic Physical Concepts Discovery with Object-Centric Predictive Models. 23252-23261
- Yatai Ji, Junjie Wang, Yuan Gong, Lin Zhang, Yanru Zhu, Hongfa Wang, Jiaxing Zhang, Tetsuya Sakai, Yujiu Yang :
MAP: Multimodal Uncertainty-Aware Vision-Language Pre-training Model. 23262-23271
- Xu Zhang, Wen Wang, Zhe Chen , Yufei Xu, Jing Zhang, Dacheng Tao:
CLAMP: Prompt-based Contrastive Learning for Connecting Language and Animal Pose. 23272-23281
- Yushi Yao, Chang Ye, Junfeng He, Gamaleldin F. Elsayed:
Teacher-generated spatial-attention labels boost robustness and accuracy of contrastive models. 23282-23291
- Yihao Liu , Jingwen He, Jinjin Gu, Xiangtao Kong, Yu Qiao, Chao Dong:
DegAE: A New Pretraining Paradigm for Low-Level Vision. 23292-23303
- Shusheng Yang, Yixiao Ge, Kun Yi, Dian Li, Ying Shan, Xiaohu Qie, Xinggang Wang:
RILS: Masked Visual Reconstruction in Language Semantic Space. 23304-23314
- Hyundo Lee, Inwoo Hwang, Hyunsung Go, Won-Seok Choi, Kibeom Kim, Byoung-Tak Zhang:
Learning Geometry-aware Representations by Sketching. 23315-23326
- Zhiyu Qu, Yulia Gryaditskaya, Ke Li, Kaiyue Pang, Tao Xiang, Yi-Zhe Song:
SketchXAI: A First Look at Explainability for Human Sketches. 23327-23337
- Sungwoong Kim, Daejin Jo, Donghoon Lee, Jongmin Kim:
MAGVLT: Masked Generative Vision-and-Language Transformer. 23338-23348
- Fengyin Lin, Mingkang Li, Da Li, Timothy M. Hospedales, Yi-Zhe Song, Yonggang Qi:
Zero-Shot Everything Sketch-Based Image Retrieval, and in Explainable Style. 23349-23358
- Jianjie Luo, Yehao Li, Yingwei Pan, Ting Yao, Jianlin Feng, Hongyang Chao, Tao Mei:
Semantic-Conditional Diffusion Networks for Image Captioning. 23359-23368
- Ziniu Hu, Ahmet Iscen, Chen Sun, Zirui Wang, Kai-Wei Chang, Yizhou Sun, Cordelia Schmid, David A. Ross, Alireza Fathi:
Reveal: Retrieval-Augmented Visual-Language Pre-Training with Multi-Source Multimodal Knowledge Memory. 23369-23379
- Minsoo Kang, Doyup Lee, Jiseob Kim, Saehoon Kim, Bohyung Han:
Variational Distribution Learning for Unsupervised Text-to-Image Generation. 23380-23389
- Yanghao Li, Haoqi Fan, Ronghang Hu, Christoph Feichtenhofer, Kaiming He:
Scaling Language-Image Pre-Training via Masking. 23390-23400
- Jihye Park, Sunwoo Kim, Soohyun Kim, Seokju Cho, Jaejun Yoo, Youngjung Uh, Seungryong Kim:
LANIT: Language-Driven Image-to-Image Translation for Unlabeled Data. 23401-23411
- Seongwon Lee, Suhyeon Lee , Hongje Seong, Euntai Kim:
Revisiting Self-Similarity: Structural Embedding for Image Retrieval. 23412-23421
- Dongwon Kim, Namyup Kim, Suha Kwak:
Improving Cross-Modal Retrieval with Set of Diverse Embeddings. 23422-23431
- Floris Weers, Vaishaal Shankar, Angelos Katharopoulos, Yinfei Yang, Tom Gunter:
Masked Autoencoding Does Not Help Natural Language Supervision at Scale. 23432-23444
- Runqi Wang, Hao Zheng, Xiaoyue Duan, Jianzhuang Liu, Yuning Lu, Tian Wang, Songcen Xu, Baochang Zhang:
Few-Shot Learning with Visual Distribution Calibration and Cross-Modal Distribution Alignment. 23445-23454
- Liangdao Wang, Yan Pan, Cong Liu, Hanjiang Lai, Jian Yin, Ye Liu:
Deep Hashing with Minimal-Distance-Separated Hash Centers. 23455-23464
- Zequn Zeng, Hao Zhang, Ruiying Lu, Dongsheng Wang, Bo Chen, Zhengjue Wang:
ConZIC: Controllable Zero-shot Image Captioning by Sampling-Based Polishing. 23465-23476
- Sarah Parisot, Yongxin Yang, Steven McDonagh :
Learning to Name Classes for Vision and Language Models. 23477-23486
- Maria Leyva-Vallina, Nicola Strisciuglio, Nicolai Petkov:
Data-Efficient Large Scale Place Recognition with Graded Similarity Supervision. 23487-23496
- Lewei Yao, Jianhua Han, Xiaodan Liang, Dan Xu, Wei Zhang, Zhenguo Li, Hang Xu:
DetCLIPv2: Scalable Open-Vocabulary Object Detection Pre-training via Word-Region Alignment. 23497-23506
- Shan Ning, Longtian Qiu, Yongfei Liu, Xuming He:
HOICLIP: Efficient Knowledge Transfer for HOI Detection with Vision-Language Models. 23507-23517
- Keyan Chen, Xiaolong Jiang, Yao Hu, Xu Tang, Yan Gao, Jianqi Chen, Weidi Xie:
OvarNet: Towards Open-Vocabulary Object Attribute Recognition. 23518-23527
- Benran Hu, Junkai Huang, Yichen Liu, Yu-Wing Tai, Chi-Keung Tang:
NeRF-RPN: A general framework for object detection in NeRFs. 23528-23538
- Vibashan VS, Ning Yu, Chen Xing, Can Qin, Mingfei Gao, Juan Carlos Niebles, Vishal M. Patel, Ran Xu:
Mask-Free OVIS: Open-Vocabulary Instance Segmentation without Manual Mask Annotations. 23539-23549



- Zhenyu Xie, Zaiyu Huang, Xin Dong, Fuwei Zhao, Haoye Dong, Xijin Zhang, Feida Zhu, Xiaodan Liang:
GP-VTON: Towards General Purpose Virtual Try-On via Collaborative Local-Flow Global-Parsing Learning. 23550-23559
- Xiaocheng Lu, Song Guo, Ziming Liu, Jingcai Guo:
Decomposed Soft Prompt Guided Fusion Enhancing for Compositional Zero-Shot Learning. 23560-23569
- Jiajin Tang , Ge Zheng , Cheng Shi , Sibei Yang:
Contrastive Grouping with Transformer for Referring Image Segmentation. 23570-23580
- (Withdrawn) Semantic Prompt for Few-Shot Image Recognition. 23581-23591
- Chang Liu, Henghui Ding , Xudong Jiang :
GRES: Generalized Referring Expression Segmentation. 23592-23601
- Qianli Feng, Raghudeep Gade, Wentong Liao, Eduard Ramon, Aleix Martinez:
Network-Free, Unsupervised Semantic Segmentation with Synthetic Images. 23602-23610
- Marlène Careil, Jakob Verbeek, Stéphane Lathuilière:
Few-shot Semantic Image Synthesis with Class Affinity Transfer. 23611-23620
- Deyi Ji, Feng Zhao, Hongtao Lu, Mingyuan Tao, Jieping Ye:
Ultra-High Resolution Segmentation with Ultra-Rich Context: A Novel Benchmark. 23621-23630
- Chenyang Lu , Daan de Geus , Gijs Dubbelman:
Content-aware Token Sharing for Efficient Semantic Segmentation with Vision Transformers. 23631-23640
- Bohao Peng, Zhuotao Tian, Xiaoyang Wu, Chengyao Wang, Shu Liu, Jingyong Su, Jiaya Jia:
Hierarchical Dense Correlation Distillation for Few-Shot Segmentation. 23641-23651
- Dongdong Wang , Boqing Gong, Liqiang Wang:
On Calibrating Semantic Segmentation Models: Analyses and An Algorithm. 23652-23662
- Junjie He, Pengyu Li, Yifeng Geng, Xuansong Xie:
FastInst: A Simple Query-Based Model for Real-Time Instance Segmentation. 23663-23672
- Zesen Cheng, Pengchong Qiao, Kehan Li, Siheng Li, Pengxu Wei, Xiangyang Ji, Li Yuan, Chang Liu, Jie Chen:
Out-of-Candidate Rectification for Weakly Supervised Semantic Segmentation. 23673-23684
- Chaohui Yu, Qiang Zhou, Jingliang Li, Jianlong Yuan, Zhibin Wang, Fan Wang:
Foundation Model Drives Weakly Incremental Learning for Semantic Segmentation. 23685-23694
- Yan Jin, Mengke Li, Yang Lu, Yiu-ming Cheung, Hanzi Wang:
Long-Tailed Visual Recognition via Self-Heterogeneous Integration with Knowledge Excavation. 23695-23704
- Zhen Zhao , Sifan Long , Jimin Pi, Jingdong Wang, Luping Zhou:
Instance-Specific and Model-Adaptive Supervision for Semi-Supervised Semantic Segmentation. 23705-23714
- Yichen Xie, Han Lu, Junchi Yan, Xiaokang Yang, Masayoshi Tomizuka, Wei Zhan:
Active Finetuning: Exploiting Annotation Budget in the Pretraining-Finetuning Paradigm. 23715-23724
- Ruo Yang, Binghui Wang, Mustafa Bilgic:
IDGI: A Framework to Eliminate Explanatory Noise from Integrated Gradients. 23725-23734
- Zhenchao Tang, Hualin Yang, Calvin Yu-Chian Chen:
Weakly Supervised Posture Mining for Fine-Grained Classification. 23735-23744
- Shiyi Lan, Xitong Yang, Zhiding Yu, Zuxuan Wu, José M. Álvarez, Anima Anandkumar:
Vision Transformers are Good Mask Auto-Labelers. 23745-23755
- Fangyi Chen, Han Zhang, Kai Hu, Yu-Kai Huang, Chenchen Zhu, Marios Savvides:
Enhanced Training of Query-Based Object Detection via Selective Query Recollection. 23756-23765
- Mengyao Lyu, Jundong Zhou, Hui Chen, Yijie Huang, Dongdong Yu, Yaqian Li, Yandong Guo, Yuchen Guo, Liuyu Xiang, Guiguang Ding:
Box-Level Active Detection. 23766-23775
- Yabo Liu, Jinghua Wang, Chao Huang, Yaowei Wang, Yong Xu:
CIGAR: Cross-Modality Graph Reasoning for Domain Adaptive Object Detection. 23776-23786
- Jingyi Zhang, Jiaxing Huang, Zhipeng Luo, Gongjie Zhang, Xiaoqin Zhang, Shijian Lu:
DA-DETR: Domain Adaptive Detection Transformer with Information Fusion. 23787-23798
- Yaoyao Liu , Bernt Schiele, Andrea Vedaldi, Christian Rupprecht:
Continual Detection Transformer for Incremental Object Detection. 23799-23808
- Jiacheng Zhang, Xiangru Lin, Wei Zhang, Kuo Wang, Xiao Tan, Junyu Han, Errui Ding, Jingdong Wang, Guanbin Li:
Semi-DETR: Semi-Supervised Object Detection with Detection Transformers. 23809-23818
- Chuandong Liu, Chenqiang Gao, Fangcen Liu, Pengcheng Li, Deyu Meng, Xinbo Gao:
Hierarchical Supervision and Shuffle Data Augmentation for 3D Semi-Supervised Object Detection. 23819-23828
- Jinhong Deng, Dongli Xu, Wen Li, Lixin Duan:
Harmonious Teacher for Cross-Domain Object Detection. 23829-23838
- Shengcao Cao, Dhiraj Joshi, Liang-Yan Gui, Yu-Xiong Wang:
Contrastive Mean Teacher for Domain Adaptive Object Detectors. 23839-23848



- Yu Wang, Pengchong Qiao, Chang Liu, Guoli Song, Xiaowu Zheng, Jie Chen:
Out-of-Distributed Semantic Pruning for Robust Semi-Supervised Learning. 23849-23858
- Ziming Liu, Song Guo, Xiaocheng Lu, Jingcai Guo, Jiewei Zhang, Yue Zeng, Fushuo Hu :
(ML)2P-Encoder: On Exploration of Channel-Class Correlation for Multi-Label Zero-Shot Learning. 23859-23868
- Duowen Chen, Yunhao Bai, Wei Shen, Qingli Li, Lequan Yu , Yan Wang:
MagicNet: Semi-Supervised Multi-Organ Segmentation via Magic-Cube Partition and Recovery. 23869-23878
- Mingze Yuan, Yingda Xia, Hexin Dong, Zifan Chen, Jiawen Yao, Mingyan Qiu, Ke Yan, Xiaoli Yin, Yu Shi , Xin Chen, Zaiyi Liu, Bin Dong, Jingren Zhou, Le Lu, Ling Zhang, Li Zhang:
Devil is in the Queries: Advancing Mask Transformers for Real-world Medical Image Segmentation and Out-of-Distribution Localization. 23879-23889
- Tiange Xiang, Yixiao Zhang, Yongyi Lu, Alan L. Yuille, Chaoyi Zhang, Weidong Cai, Zongwei Zhou:
SQUID: Deep Feature In-Painting for Unsupervised Anomaly Detection. 23890-23901
- Jeongun Ryu, Aaron Valero Puche, Jaewoong Shin, Seonwook Park, Biagio Brattoli, Jinhee Lee, Wonkyung Jung, Soo Ick Cho, Kyunghyun Paeng, Chan-Young Ock, Donggeun Yoo, Sérgio Pereira :
OCELOT: Overlapped Cell on Tissue Dataset for Histopathology. 23902-23912
- Aayush Kumar Tyagi, Chirag Mohapatra, Prasenjit Das, Govind Makharia, Lalita Mehra, Prathosh AP, Mausam:
DeGPR: Deep Guided Posterior Regularization for Multi-Class Cell Detection and Counting. 23913-23923
- Paul Hager, Martin J. Menten, Daniel Rueckert:
Best of Both Worlds: Multimodal Contrastive Learning with Tabular and Imaging Data. 23924-23935
- Yuan-Chih Chen, Chun-Shien Lu:
RankMix: Data Augmentation for Weakly Supervised Learning of Classifying Whole Slide Images with Diverse Sizes and Imbalanced Categories. 23936-23945
- Xixi Liu, Yaroslava Lochman, Christopher Zach:
GEN: Pushing the Limits of Softmax-Based Out-of-Distribution Detection. 23946-23955
- Aming Wu, Cheng Deng:
Discriminating Known from Unknown Objects via Structure-Enhanced Recurrent Variational AutoEncoder. 23956-23965
- Yuze Tan, Yixi Liu, Shudong Huang, Wentao Feng, Jiancheng Lv:
Sample-level Multi-view Graph Clustering. 23966-23975
- Daniel J. Trosten , Sigurd Løkse, Robert Jenssen, Michael C. Kampffmeyer:
On the Effects of Self-supervision and Contrastive Alignment in Deep Multi-view Clustering. 23976-23985
- Pengxin Zeng, Yunfan Li, Peng Hu, Dezhong Peng, Jiancheng Lv, Xi Peng:
Deep Fair Clustering via Maximizing and Minimizing Mutual Information: Theory, Algorithm and Metric. 23986-23995
- Hao Zhu, Piotr Koniusz:
Transductive Few-Shot Learning with Prototype-Based Label Propagation by Iterative Graph Refinement. 23996-24006
- Malik Boudiaf, Etienne Bennequin, Myriam Tami, Antoine Toubhans, Pablo Piantanida, Céline Hudelot, Ismail Ben Ayed:
Open-Set Likelihood Maximization for Few-Shot Learning. 24007-24016
- Beitong Zhou, Jing Lu, Kerui Liu, Yunlu Xu, Zhanzhan Cheng, Yi Niu:
HyperMatch: Noise-Tolerant Semi-Supervised Learning via Relaxed Contrastive Constraint. 24017-24026
- Tianjiao Li, Lin Geng Foo, Ping Hu, Xindi Shang, Hossein Rahmani, Zehuan Yuan, Jun Liu:
Token Boosting for Robust Self-Supervised Visual Transformer Pre-training. 24027-24038
- Taeuk Jang, Xiaoqian Wang:
Difficulty-Based Sampling for Debiased Contrastive Representation Learning. 24039-24048
- Corentin Dancette, Spencer Whitehead, Rishabh Maheshwary, Ramakrishna Vedantam, Stefan Scherer, Xinlei Chen, Matthieu Cord, Marcus Rohrbach:
Improving Selective Visual Question Answering by Learning from Your Peers. 24049-24059
- Zeyu Gan, Suyun Zhao, Jinlong Kang, Liyuan Shang, Hong Chen, Cuiping Li:
Superclass Learning with Representation Enhancement. 24060-24069
- Yifan Li, Hu Han, Shiguang Shan, Xilin Chen:
DISC: Learning from Noisy Labels via Dynamic Instance-Specific Selection and Correction. 24070-24079
- Jian Li, Ziyao Meng, Daqian Shi, Rui Song, Xiaolei Diao, Jingwen Wang, Hao Xu:
FCC: Feature Clusters Compression for Long-Tailed Visual Recognition. 24080-24089
- Wei Wang, Zhun Zhong, Weijie Wang, Xi Chen, Charles Ling, Boyu Wang, Nicu Sebe :
Dynamically Instance-Guided Adaptation: A Backward-free Approach for Test-Time Domain Adaptive Semantic Segmentation. 24090-24099
- Yu-Chu Yu, Hsuan-Tien Lin:
Semi-Supervised Domain Adaptation with Source Label Adaptation. 24100-24109
- Wuyang Li , Jie Liu , Bo Han, Yixuan Yuan :
Adjustment and Alignment for Unbiased Open Set Domain Adaptation. 24110-24119



- Nazmul Karim, Niluthpol Chowdhury Mithun, Abhinav Rajvanshi, Han-Pang Chiu, Supun Samarasekera, Nazanin Rahnavard:
C-SFDA: A Curriculum Learning Aided Self-Training Framework for Efficient Source Free Domain Adaptation. 24120-24131
- Jintao Guo, Na Wang, Lei Qi, Yinghuan Shi:
ALOFT: A Lightweight MLP-Like Architecture with Dynamic Low-Frequency Transform for Domain Generalization. 24132-24141
- Sanqing Qu, Yingwei Pan, Guang Chen, Ting Yao, Changjun Jiang, Tao Mei:
Modality-Agnostic Debiasing for Single Domain Generalization. 24142-24151
- Muhammad Jehanzeb Mirza, Pol Jané-Soneira, Wei Lin, Mateusz Kozinski, Horst Possegger, Horst Bischof :
ActMAD: Activation Matching to Align Distributions for Test-Time-Training. 24152-24161
- A. Tuan Nguyen, Thanh Nguyen-Tang, Ser-Nam Lim, Philip H. S. Torr:
TIPI: Test Time Adaptation with Transformation Invariance. 24162-24171
- Liang Chen, Yong Zhang, Yibing Song, Ying Shan, Lingqiao Liu:
Improved Test-Time Adaptation for Domain Generalization. 24172-24182
- Zeyin Song, Yifan Zhao, Yujun Shi, Peixi Peng, Li Yuan, Yonghong Tian:
Learning with Fantasy: Semantic-Aware Virtual Contrastive Constraint for Few-Shot Class-Incremental Learning. 24183-24192
- Karim Guirguis, Johannes Meier, George Eskandar, Matthias Kayser, Bin Yang, Jürgen Beyerer:
NIFF: Alleviating Forgetting in Generalized Few-Shot Object Detection via Neural Instance Feature Forging. 24193-24202
- Jingjing Jiang, Nanning Zheng:
MixPHM: Redundancy-Aware Parameter-Efficient Tuning for Low-Resource Visual Question Answering. 24203-24213
- Andrés Villa, Juan León Alcázar, Motasem Alfarrà, Kumail Alhamoud, Julio Hurtado, Fabian Caba Heilbron, Alvaro Soto, Bernard Ghanem :
PIVOT: Prompting for Video Continual Learning. 24214-24223
- Changdae Oh, Hyeji Hwang, Hee Young Lee, YongTaek Lim, Geunyoung Jung, Jiyoung Jung, Hosik Choi, Kyungwoo Song:
BlackVIP: Black-Box Visual Prompting for Robust Transfer Learning. 24224-24235
- Xinyuan Gao, Yuhang He, Songlin Dong, Jie Cheng, Xing Wei, Yihong Gong:
DKT: Diverse Knowledge Transfer Transformer for Class Incremental Learning. 24236-24245
- Huiwei Lin, Baoquan Zhang, Shanshan Feng, Xutao Li, Yunming Ye:
PCR: Proxy-Based Contrastive Replay for Online Class-Incremental Continual Learning. 24246-24255
- Yutong Bai, Zeyu Wang, Junfei Xiao, Chen Wei, Huiyu Wang, Alan L. Yuille, Yuyin Zhou, Cihang Xie:
Masked Autoencoders Enable Efficient Knowledge Distillers. 24256-24265
- Shikang Yu, Jiachen Chen, Hu Han, Shuqiang Jiang:
Data-Free Knowledge Distillation via Feature Exchange and Activation Region Constraint. 24266-24275
- Ying Jin, Jiaqi Wang, Dahua Lin:
Multi-Level Logit Distillation. 24276-24285
- Qiao Gu, Dongsub Shim, Florian Shkurti:
Preserving Linear Separability in Continual Learning by Backward Feature Projection. 24286-24295
- Michael Kleinman, Alessandro Achille, Stefano Soatto:
Critical Learning Periods for Multisensory Integration in Deep Networks. 24296-24305
- Juliette Marrie, Michael Arbel, Diane Larlus, Julien Mairal:
SLACK: Stable Learning of Augmentations with Cold-Start and KL Regularization. 24306-24314
- Fangrui Lv, Jian Liang, Shuang Li, Jinming Zhang, Di Liu:
Improving Generalization with Domain Convex Game. 24315-24324
- Zhi Gao, Chen Xu , Feng Li, Yunde Jia, Mehrtash Harandi, Yuwei Wu:
Exploring Data Geometry for Continual Learning. 24325-24334
- Xingchao Liu, Lemeng Wu, Shujian Zhang, Chengyue Gong, Wei Ping, Qiang Liu:
FlowGrad: Controlling the Output of Generative ODEs with Gradients. 24335-24344
- Yongcheng Jing, Chongbin Yuan, Li Ju, Yiding Yang, Xinchao Wang, Dacheng Tao:
Deep Graph Reprogramming. 24345-24354
- Lu Yu, Wei Xiang:
X-Pruner: eXplainable Pruning for Vision Transformers. 24355-24363
- Eugenia Iofinova, Alexandra Peste, Dan Alistarh:
Bias in Pruned Vision Models: In-Depth Analysis and Countermeasures. 24364-24373
- Yikai Wang, Wenbing Huang, Yinpeng Dong, Fuchun Sun, Anbang Yao:
Compacting Binary Neural Networks by Sparse Kernel Selection. 24374-24383
- Jishnu Mukhoti, Andreas Kirsch, Joost van Amersfoort, Philip H. S. Torr, Yarin Gal:
Deep Deterministic Uncertainty: A New Simple Baseline. 24384-24394



- Suman V. Ravuri, Mélanie Rey, Shakir Mohamed, Marc Peter Deisenroth:
Understanding Deep Generative Models with Generalized Empirical Likelihoods. 24395-24405
- Pengwei Tang, Wei Yao, Zhicong Li, Yong Liu:
Fair Scratch Tickets: Finding Fair Sparse Networks without Weight Training. 24406-24416
- Huantong Li, Xiangmiao Wu, Fanbing Lv, Daihai Liao, Thomas H. Li, Yonggang Zhang, Bo Han, Mingkui Tan:
Hard Sample Matters a Lot in Zero-Shot Quantization. 24417-24426
- Jiawei Liu, Lin Niu, Zhihang Yuan , Dawei Yang, Xinggang Wang, Wenyu Liu:
PD-Quant: Post-Training Quantization Based on Prediction Difference Metric. 24427-24437
- Zhou Yang, Weisheng Dong, Xin Li, Mengluan Huang, Yulin Sun, Guangming Shi:
Vector Quantization with Self-Attention for Quality-Independent Representation Learning. 24438-24448
- Zhengcong Fei, Mingyuan Fan, Li Zhu, Junshi Huang, Xiaoming Wei, Xiaolin Wei:
Masked Auto-Encoders Meet Generative Adversarial Networks and Beyond. 24449-24459
- Arkanath Pathak, Nicholas Dufour:
Sequential Training of GANs Against GAN-Classifiers Reveals Correlated "Knowledge Gaps" Present Among Independently Trained GAN Instances. 24460-24469
- Aditay Tripathi, Rishabh Singh, Anirban Chakraborty, Pradeep Shenoy:
Edges to Shapes to Concepts: Adversarial Augmentation for Robust Vision. 24470-24479
- Utkarsh Ojha, Yuheng Li, Yong Jae Lee:
Towards Universal Fake Image Detectors that Generalize Across Generative Models. 24480-24489
- Xincheng Yao, Ruoqi Li, Jing Zhang, Jun Sun, Chongyang Zhang:
Explicit Boundary Guided Semi-Push-Pull Contrastive Learning for Supervised Anomaly Detection. 24490-24499
- Zuhao Liu, Xiao-Ming Wu, Dian Zheng, Kun-Yu Lin, Wei-Shi Zheng:
Generating Anomalies for Video Anomaly Detection with Prompt-based Feature Mapping. 24500-24510
- Tran Dinh Tien, Anh Tuan Nguyen, Nguyen Hoang Tran, Ta Duc Huy, Soan Thi Minh Duong, Chanh D. Tr. Nguyen, Steven Q. H. Truong:
Revisiting Reverse Distillation for Anomaly Detection. 24511-24520
- Zhenyi Wang, Li Shen, Donglin Zhan, Qiuling Suo, Yanjun Zhu, Tiehang Duan, Mingchen Gao:
MetaMix: Towards Corruption-Robust Continual Learning with Temporally Self-Adaptive Data Transformation. 24521-24531
- Fatih Ilhan , Gong Su, Ling Liu:
ScaleFL: Resource-Adaptive Federated Learning with Heterogeneous Clients. 24532-24541
- Junyi Zhu, Xingchen Ma, Matthew B. Blaschko :
Confidence-Aware Personalized Federated Learning via Variational Expectation Maximization. 24542-24551
- Yifan Shi, Yingqi Liu, Kang Wei, Li Shen, Xueqian Wang, Dacheng Tao:
Make Landscape Flatter in Differentially Private Federated Learning. 24552-24562
- Yiyou Sun, Yaojie Liu, Xiaoming Liu, Yixuan Li, Wen-Sheng Chu:
Rethinking Domain Generalization for Face Anti-spoofing: Separability and Alignment. 24563-24574
- Yuqian Fu, Yu Xie, Yanwei Fu , Yu-Gang Jiang:
StyleAdv: Meta Style Adversarial Training for Cross-Domain Few-Shot Learning. 24575-24584
- Simin Chen, Hanlin Chen, Mirazul Haque, Cong Liu, Wei Yang:
The Dark Side of Dynamic Routing Neural Networks: Towards Efficiency Backdoor Injection. 24585-24594
- Mikel Bober-Irizar, Ilia Shumailov, Yiren Zhao, Robert D. Mullins, Nicolas Papernot:
Architectural Backdoors in Neural Networks. 24595-24604
- Zenghui Yuan, Pan Zhou, Kai Zou, Yu Cheng:
You Are Catching My Attention: Are Vision Transformers Bad Learners under Backdoor Attacks? 24605-24615
- Fan Wang, Adams Wai-Kin Kong:
A Practical Upper Bound for the Worst-Case Attribution Deviations. 24616-24625
- Zexin Li, Bangjie Yin, Taiping Yao, Junfeng Guo, Shouhong Ding, Simin Chen, Cong Liu:
Sibling-Attack: Rethinking Transferable Adversarial Attacks against Face Recognition. 24626-24637
- Wenwen Si, Shuo Li, Sangdon Park, Insup Lee, Osbert Bastani:
Angelic Patches for Improving Third-Party Object Detector Performance. 24638-24647
- Junyoung Byun, Myung-Joon Kwon, Seungju Cho, Yoonji Kim, Changick Kim:
Introducing Competition to Boost the Transferability of Targeted Adversarial Examples Through Clean Feature Mixup. 24648-24657
- Lei Hsiung, Yun-Yun Tsai, Pin-Yu Chen, Tsung-Yi Ho:
Towards Compositional Adversarial Robustness: Generalizing Adversarial Training to Composite Semantic Perturbations. 24658-24667
- Bo Huang, Mingyang Chen, Yi Wang, Junda Lu, Minhao Cheng , Wei Wang:
Boosting Accuracy and Robustness of Student Models via Adaptive Adversarial Distillation. 24668-24677
- Junhao Dong, Seyed-Mohsen Moosavi-Dezfooli, Jianhuang Lai, Xiaohua Xie:
The Enemy of My Enemy is My Friend: Exploring Inverse Adversaries for Improving Adversarial Training. 24678-24687



24687

- Zhenbo Song, Zhenyuan Zhang , Kaihao Zhang, Wenhan Luo , Zhaoxin Fan, Wenqi Ren, Jianfeng Lu:
Robust Single Image Reflection Removal Against Adversarial Attacks. 24688-24698
- Yanjie Li, Yiquan Li, Xuelong Dai, Songtao Guo, Bin Xiao:
Physical-World Optical Adversarial Attacks on 3D Face Recognition. 24699-24708
- Weiming Bai, Yufan Liu, Zhipeng Zhang, Bing Li, Weiming Hu:
AUNet: Learning Relations Between Action Units for Face Forgery Detection. 24709-24719

VALID HTML last updated on 2024-12-11 20:32 CET by the dblp team

all metadata released as open data under CC0 1.0 license

see also: Terms of Use | Privacy Policy | Imprint

dblp was originally created in 1993 at:



since 2018, dblp has been operated and maintained by:



the dblp computer science bibliography is funded and supported by:

