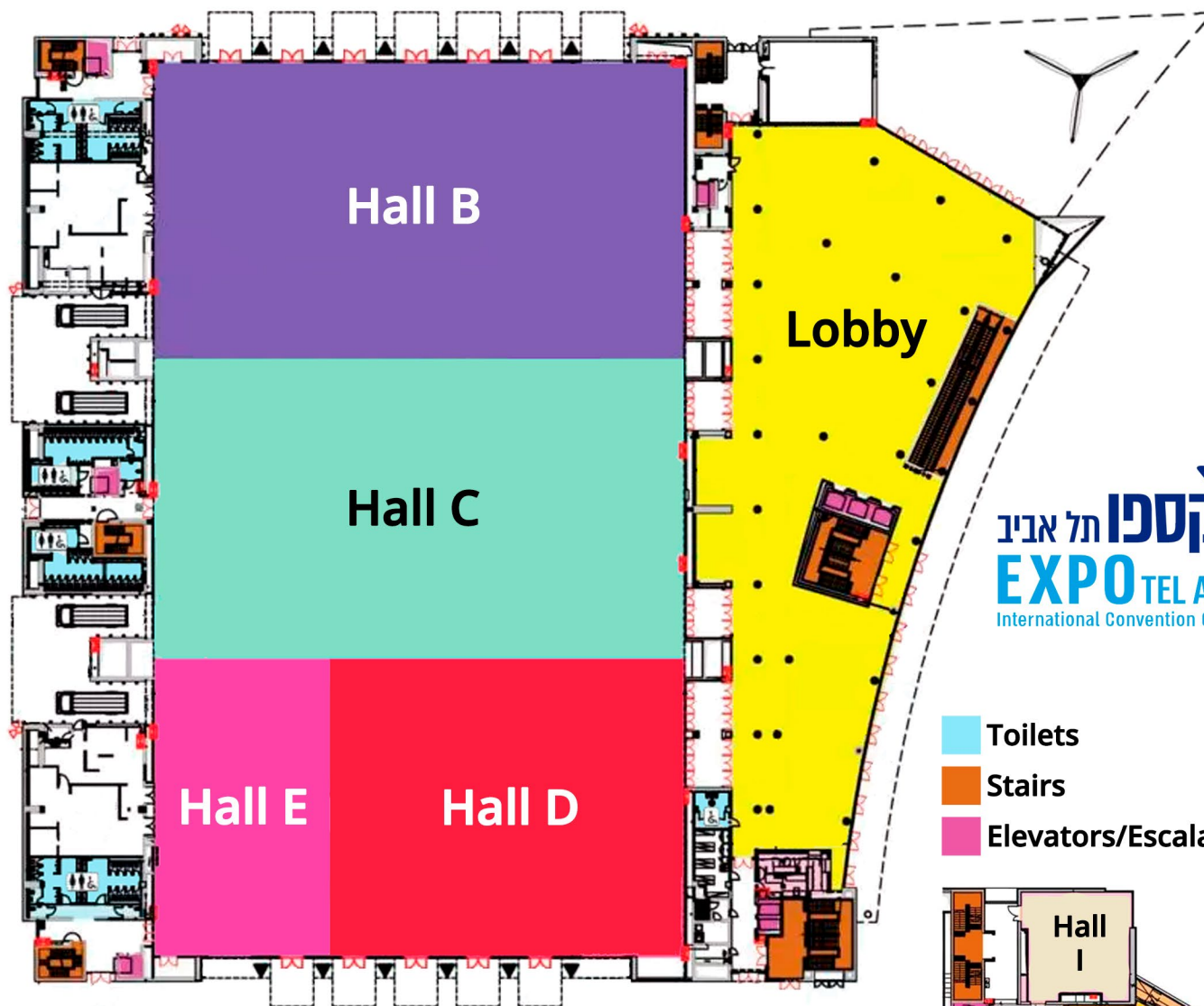




EUROPEAN CONFERENCE
ON COMPUTER VISION
TEL AVIV 2022

Program Guide
Main Conference

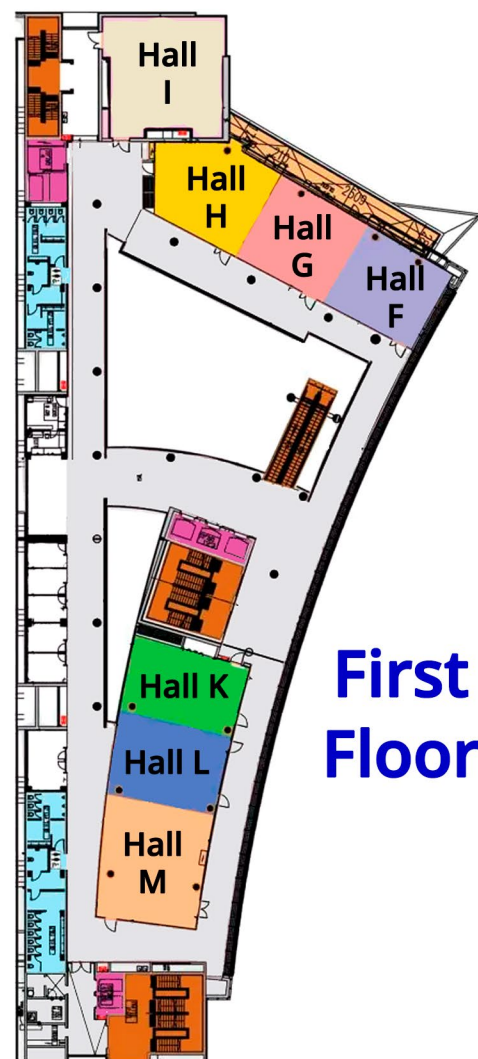




Ground Floor

Tel Aviv
Convention Center
Pavilion 2

- Toilets
- Stairs
- Elevators/Escalator



First Floor

Message from the General and Program Chairs

Welcome to the 2022 European Conference on Computer Vision in Tel-Aviv, Israel. This is the place to see the latest research results, listen to distinguished keynote speakers, attend tutorials, workshops and demos, and have a chance to meet old friends and make new ones.

The proceedings you virtually hold in your hands are the result of a collective effort. 18310 authors submitted a total of 6773 papers that were handled by 276 area chairs (ACs) that solicited the help of 4719 reviewers. The entire process was supervised by the program chairs (PCs) with the constant support of the general chairs (GCs). In the end, we have 1645 (28%) papers that were accepted for publication, including 157 Orals (2.7%).

The double-blind reviewing process was handled by the CMT system. Authors did not know the name of the reviewers and vice versa. 846 of the submissions were desk-rejected for various reasons. Many of them because they revealed author identity, thus violating the double-blind policy. Some papers were withdrawn at different stages leaving us with a total 5804 valid submissions.

Each of these submissions received at least three reviews (except 6 papers that received only 2 reviews), totalling more than 15,000 reviews. Authors had a chance to submit a rebuttal, followed by a discussion between the area chairs (ACs) and the reviewers assigned to each paper. The final decision of each paper was taken by the AC in consultation with a buddy AC, to make sure the decisions are as fair and informative as possible. The process was monitored by the PCs, with a special emphasis on cases where the decision of the AC differs from the consensus recommendation of the reviewers.

The planning of ECCV 2022 had to deal with the uncertainties of the COVID-19 pandemic. ECCV 2022 is, still, a hybrid conference that gives researchers the ability to attend the conference either virtually or in-person. Based on past experience, we have decided that only in-person attendees can present their work on site. In addition, all attendees, in-person or virtual, can watch a 5-minute video of each of the papers on the virtual platform.

The conference runs for three days and includes two parallel oral sessions per day, as well as two poster sessions per day. In addition, there are two days of workshops and tutorials, as well as events that happen in parallel to the main conference, including mentoring sessions, industrial exhibition, academic demos, and an industrial track focused on entrepreneurs.

A separate committee was tasked with selecting the best paper award, along with the honorable mentions. The selected papers will be announced in a special session during the conference.

We have a long list of people to thank. We thank Pavel Lifshitz, our Technical Program Chair, for working tirelessly behind the scenes. We thank our demo chairs, workshop chairs, communication chairs, best-paper committee members, social activities chairs, industry track chairs, and the diversity committee members for helping us along the way. We benefited from the experience and advice of Nicole Finn regarding organizational aspects and thank her for that. A special thanks to the entire ORTRA team for organizing the conference. And last, but not least, we thank you for submitting a paper, reviewing papers and attending. We hope you enjoy ECCV'2022!

Program Chairs: Shai Avidan, Gabriel Brostow, Giovanni Maria Farinella, and Tal Hassner

General Chairs: Rita Cucchiara, Jiří Matas, Amnon Shashua, and Lihi Zelnik-Manor

ECCV 2022 Organizing Committee

General Chairs: Rita Cucchiara
Jiří Matas
Amnon Shashua
Lihi Zelnik-Manor

Program Chairs: Shai Avidan
Gabriel Brostow
Giovanni Maria Farinella
Tal Hassner

Program Technical Chair: Pavel Lifshitz

Workshops Chairs: Leonid Karlinsky
Tomer Michaeli
Ko Nishino

Tutorials Chairs: Thomas Pock
Natalia Neverova

Demo Chair: Bohyung Han

Finance Chair: Gerard Medioni

Publications Chair: Eric Mortensen

Social & Student Activities Chairs: Tatiana Tommasi
Sagie Benaïm

Industrial Liaison Chairs: Dimosthenis Karatzas
Chen Sagiv

Industry Track Chairs: Amir Markovitz
Yair Kittenplon

Communications Chairs: Lorenzo Baraldi
Kosta Derpanis

Diversity and Inclusion Chairs: Xi Yin
Bryan Russell

Award Papers Committee: Laura Leal-Taixé (chair)
Ko Nishino
Yanxi Liu
Alessio del Bue
Todd Zickler
Jianbo Shi
Oisín Mac Aodha

ECCV 2022 Area Chairs

Lourdes Agapito	Basura Fernando	Benjamin Kimia	Cuong V. Nguyen	Yi-Zhe Song
Zeynep Akata	Vittorio Ferrari	Alexander Kirillov	Bingbing Ni	Bjorn Stenger
Naveed Akhtar	Andrew W. Fitzgibbon	Kris Kitani	Juan Carlos Niebles	Abby Stylianou
Karteek Alahari	David J. Fleet	Iasonas Kokkinos	Ko Nishino	Akihiro Sugimoto
Alexandre Alahi	David Forsyth	Vladlen Koltun	Jean-Marc Odobez	Chen Sun
Pablo Arbelaez	David Fouhey	Nikos Komodakis	Francesca Odone	Deqing Sun
Antonis A. Argyros	Katerina Fragkiadaki	Piotr Koniusz	Takayuki Okatani	Kalyan Sunkavalli
Yuki M. Asano	Friedrich Fraundorfer	Philipp Kraehenbuehl	Manohar Paluri	Ying Tai
Kalle Åström	Oren Freifeld	Dilip Krishnan	Guan Pang	Ayellet Tal
Hadar Averbuch-Elor	Thomas Funkhouser	Ajay Kumar	Maja Pantic	Ping Tan
Hossein Azizpour	Yasutaka Furukawa	Junseok Kwon	Sylvain Paris	Siyu Tang
Vineeth N. Balasubramanian	Fabio Galasso	Jean-Francois Lalonde	Jaesik Park	Chi-Keung Tang
Lamberto Ballan	Jürgen Gall	Ivan Laptev	Hyun Soo Park	Radu Timofte
Adrien Bartoli	Chuang Gan	Laura Leal-Taixé	Omkar M. Parkhi	Federico Tombari
Horst Bischof	Zhe Gan	Erik Learned-Miller	Deepak Pathak	James Tompkin
Matthew B. Blaschko	Animesh Garg	Gim Hee Lee	Georgios Pavlakos	Lorenzo Torresani
Federica Bogo	Efstratios Gavves	Seungyong Lee	Marcello Pelillo	Alexander Toshev
Katherine Bouman	Peter Gehler	Zhen Lei	Marc Pollefeys	Du Tran
Edmond Boyer	Theo Gevers	Bastian Leibe	Jean Ponce	Anh T. Tran
Michael S. Brown	Bernard Ghanem	Hongdong Li	Gerard Pons-Moll	Zhuowen Tu
Vittorio Caggiano	Ross B. Girshick	Fuxin Li	Fatih Porikli	Georgios Tzimiropoulos
Neill Campbell	Georgia Gkioxari	Bo Li	Victor Adrian Prisacariu	Jasper Uijlings
Octavia Camps	Albert Gordo	Yin Li	Petia Radeva	Jan C. van Gemert
Duygu Ceylan	Stephen Gould	Ser-Nam Lim	Ravi Ramamoorthi	Gul Varol
Ayan Chakrabarti	Venu Madhav Govindu	Joseph Lim	Deva Ramanan	Nuno Vasconcelos
Tat-Jen Cham	Kristen Grauman	Stephen Lin	Vignesh Ramanathan	Mayank Vatsa
Antoni Chan	Abhinav Gupta	Dahua Lin	Nalini Ratha	Ashok Veeraraghavan
Manmohan Chandraker	Mohit Gupta	Si Liu	Tammy Riklin Raviv	Jakob Verbeek
Xinlei Chen	Hu Han	Xiaoming Liu	Tobias Ritschel	Carl Vondrick
Xilin Chen	Bohyung Han	Ce Liu	Emanuele Rodola	Ruiping Wang
Dongdong Chen	Tian Han	Zicheng Liu	Amit K. Roy-Chowdhury	Xinchao Wang
Chen Chen	Emily Hand	Yanxi Liu	Michael Rubinstein	Liwei Wang
Ondrej Chum	Bharath Hariharan	Feng Liu	Olga Russakovsky	Chaohui Wang
John Collomosse	Ran He	Yebin Liu	Mathieu Salzmann	Xiaolong Wang
Camille Couprie	Otmar Hilliges	Chen Change Loy	Dimitris Samaras	Christian Wolf
David Crandall	Adrian Hilton	Huchuan Lu	Aswin Sankaranarayanan	Tao Xiang
Daniel Cremers	Minh Hoai	Cewu Lu	Imari Sato	Saining Xie
Marco Cristani	Yedid Hoshen	Oisin Mac Aodha	Yoichi Sato	Cihang Xie
Canton Cristian	Timothy Hospedales	Dhruv Mahajan	Shin'ichi Satoh	Zeki Yalniz
Dengxin Dai	Gang Hua	Subhransu Maji	Walter Scheirer	Ming-Hsuan Yang
Dima Damen	Di Huang	Atsuto Maki	Bernt Schiele	Angela Yao
Kostas Daniilidis	Jing Huang	Arun Mallya	Konrad Schindler	Shaodi You
Trevor Darrell	Jia-Bin Huang	R. Manmatha	Cordelia Schmid	Stella X. Yu
Andrew Davison	Nathan Jacobs	Iacopo Masi	Alexander Schwing	Junsong Yuan
Tali Dekel	C.V. Jawahar	Dimitris N. Metaxas	Nicu Sebe	Stefanos Zafeiriou
Alessio Del Bue	Herve Jegou	Ajmal Mian	Greg Shakhnarovich	Amir Zamir
Weihong Deng	Neel Joshi	Christian Micheloni	Eli Shechtman	Lei Zhang
Konstantinos Derpanis	Armand Joulin	Krystian Mikolajczyk	Humphrey Shi	Lei Zhang
Carl Doersch	Frederic Jurie	Anurag Mittal	Jianbo Shi	Pengchuan Zhang
Matthijs Douze	Fredrik Kahl	Philippos Mordohai	Roy Shilkrot	Bolei Zhou
Mohamed Elhoseiny	Yannis Kalantidis	Greg Mori	Mike Zheng Shou	Yuke Zhu
Sergio Escalera	Evangelos Kalogerakis	Vittorio Murino	Kaleem Siddiqi	Todd Zickler
Yi Fang	Sing Bing Kang	P. J. Narayanan	Richa Singh	Wangmeng Zuo
Ryan Farrell	Yosi Keller	Ram Nevatia	Greg Slabaugh	
Alireza Fathi	Margret Keuper	Natalia Neverova	Cees Snoek	
Christoph Feichtenhofer	Tae-Kyun Kim	Richard Newcombe	Yale Song	

Tuesday, October 25

0700–1830 Registration (Pavillion 2 Lobby)

0800–0900 Coffee & Refreshments (Lobby and Hall C)

0830–0930 Poster Setup (Hall B)

0900–1630 Exhibits (Hall C)

- See online for exhibitor list and map.

0900–1530 Demos (Hall C Demo Area)

- Bounding Box Estimation on iPhone, *Yu Xuan Liu (UC Berkeley / Covariant.AI, USA)*
- Data Discovery and Targeted Learning, *Suraj Kothawade (Univ. of Texas at Dallas, USA)*
- CAFDL: A Fast Real-Time Fire Detection and Localization System at Edge, *Jayeeta Mondal (TCS Research, India)*
- Visualization of Knowledge and Stories about Cultural Heritage in Augmented Reality, *Stuart James, Mateo Taiana, Matteo Toso, Alessio Del Bue (Pattern Analysis and Computer Vision (PAVIS), Istituto Italiano di Tecnologia)*

0930–1100 Oral 1.A.1: Detection, Recognition, Classification, & Localization in 2D/3D (Hall D; overflow in Halls K-M)

Papers in this session are in Poster Session 1.A
Format (9 min. presentation; 3 min. questions)

- [0930] Welcome, *ECCV 2022 Program Chairs*
- [0935] Long-Tail Detection with Effective Class-Margins, *Jang Hyun Cho, Philipp Krähenbühl*
- [0947] Multimodal Object Detection via Probabilistic Ensembling, *Yi-Ting Chen, Jinghao Shi, Zelin Ye, Christoph Mertz, Deva Ramanan, Shu Kong*
- [0959] Improving Robustness by Enhancing Weak Subnets, *Yong Guo, David Stutz, Bernt Schiele*
- [1011] Vote from the Center: 6 DoF Pose Estimation in RGB-D Images by Radial Keypoint Voting, *Yangzheng Wu, Mohsen Zand, Ali Etamad, Michael Greenspan*
- [1023] Monocular 3D Object Detection with Depth from Motion, *Tai Wang, Jiangmiao Pang, Dahua Lin*

0930–1100 Oral 1.A.2: Motion and Tracking (Hall E; overflow in Halls F-H)

Papers in this session are in Poster Session 1.A
Format (9 min. presentation; 3 min. questions)

- [0930] Welcome, *ECCV 2022 Program Chairs*
- [0935] Particle Video Revisited: Tracking through Occlusions Using Point Trajectories, *Adam W. Harley, Zhaoyuan Fang, Katerina Fragkiadaki*
- [0947] A Perturbation-Constrained Adversarial Attack for Evaluating the Robustness of Optical Flow, *Jenny Schmalfuss, Philipp Scholze, Andrés Bruhn*
- [0959] Social-SSL: Self-Supervised Cross-Sequence Representation Learning Based on Transformers for Multi-agent Trajectory Prediction, *Li-Wu Tsao, Yan-Kai Wang, Hao-Siang Lin, Hong-Han Shuai, Lai-Kuan Wong, Wen-Huang Cheng*

- [1011] Diverse Human Motion Prediction Guided by Multi-level Spatial-Temporal Anchors, *Sirui Xu, Yu-Xiong Wang, Liang-Yan Gui*
- [1023] TEMOS: Generating Diverse Human Motions from Textual Descriptions, *Mathis Petrovich, Michael J. Black, Gül Varol*
- [1035] PREF: Predictability Regularized Neural Motion Fields, *Liangchen Song, Xuan Gong, Benjamin Planche, Meng Zheng, David Doermann, Junsong Yuan, Terrence Chen, Ziyang Wu*

1100–1200 Coffee & Refreshments (Lobby and Hall C)

1100–1330 Mentoring Session (Halls K-M)

1100–1330 Poster 1.A (Hall B)

- How Severe Is Benchmark-Sensitivity in Video Self-Supervised Learning? *Fida Mohammad Thoker, Hazel Doughty, Piyush Bagad, Cees G. M. Snoek*
- Decoupled Contrastive Learning, *Chun-Hsiao Yeh, Cheng-Yao Hong, Yen-Chi Hsu, Tyng-Luh Liu, Yubei Chen, Yann LeCun*
- 3D Clothed Human Reconstruction in the Wild, *Gyeongsik Moon, Hyeongjin Nam, Takaaki Shiratori, Kyoung Mu Lee*
- MTFormer: Multi-task Learning via Transformer and Cross-Task Reasoning, *Xiaogang Xu, Hengshuang Zhao, Vibhav Vineet, Ser-Nam Lim, Antonio Torralba*
- Towards Accurate Network Quantization with Equivalent Smooth Regularizer, *Kirill Solodskikh, Vladimir Chikin, Ruslan Aydarkhanov, Dehua Song, Irina Zhelavskaya, Jiansheng Wei*
- PatchRD: Detail-Preserving Shape Completion by Learning Patch Retrieval and Deformation, *Bo Sun, Vladimir G. Kim, Noam Aigerman, Qixing Huang, Siddhartha Chaudhuri*
- Object Discovery via Contrastive Learning for Weakly Supervised Object Detection, *Jinhwan Seo, Wonho Bae, Danica J. Sutherland, Junhyug Noh, Daijin Kim*
- Federated Self-Supervised Learning for Video Understanding, *Yasar Abbas Ur Rehman, Yan Gao, Jiajun Shen, Pedro Porto Buarque de Gusmão, Nicholas Lane*
- TEMOS: Generating Diverse Human Motions from Textual Descriptions, *Mathis Petrovich, Michael J. Black, Gül Varol*
- S3C: Self-Supervised Stochastic Classifiers for Few-Shot Class-Incremental Learning, *Jayateja Kalla, Soma Biswas*
- Parameterized Temperature Scaling for Boosting the Expressive Power in Post-Hoc Uncertainty Calibration, *Christian Tomani, Daniel Cremers, Florian Buettner*
- Intelli-Paint: Towards Developing More Human-Intelligible Painting Agents, *Jaskirat Singh, Cameron Smith, Jose Echevarria, Liang Zheng*
- Expanded Adaptive Scaling Normalization for End-to-End Image Compression, *Chajin Shin, Hyeongmin Lee, Hanbin Son, Sangjin Lee, Dogyoon Lee, Sangyoun Lee*
- ARAH: Animatable Volume Rendering of Articulated Human SDFs, *Shaofei Wang, Katja Schwarz, Andreas Geiger, Siyu Tang*
- Scaling Adversarial Training to Large Perturbation Bounds, *Sravanti Addepalli, Samyak Jain, Gaurang Sriramanan, R. Venkatesh Babu*
- DualFormer: Local-Global Stratified Transformer for Efficient Video Recognition, *Yuxuan Liang, Pan Zhou, Roger Zimmermann, Shuicheng Yan*

17. WISE: Whitebox Image Stylization by Example-Based Learning, *Winfried Löttsch, Max Reimann, Martin Büsselmeier, Amir Semmo, Jürgen Döllner, Matthias Trapp*
18. CramNet: Camera-Radar Fusion with Ray-Constrained Cross-Attention for Robust 3D Object Detection, *Jyh-Jing Hwang, Henrik Kretzschmar, Joshua Manela, Sean Rafferty, Nicholas Armstrong-Crews, Tiffany Chen, Dragomir Anguelov*
19. Particle Video Revisited: Tracking through Occlusions Using Point Trajectories, *Adam W. Harley, Zhaoyuan Fang, Katerina Fragkiadaki*
20. Image Inpainting with Cascaded Modulation GAN and Object-Aware Training, *Haitian Zheng, Zhe Lin, Jingwan Lu, Scott Cohen, Eli Shechtman, Connelly Barnes, Jianming Zhang, Ning Xu, Sohrab Amirghodsi, Jiebo Luo*
21. Learning to Detect Every Thing in an Open World, *Kuniaki Saito, Ping Hu, Trevor Darrell, Kate Saenko*
22. Backbone Is All Your Need: A Simplified Architecture for Visual Object Tracking, *Boyu Chen, Peixia Li, Lei Bai, Lei Qiao, Qihong Shen, Bo Li, Weihao Gan, Wei Wu, Wanli Ouyang*
23. Neural Correspondence Field for Object Pose Estimation, *Lin Huang, Tomas Hodan, Lingni Ma, Linguang Zhang, Luan Tran, Christopher D. Twigg, Po-Chen Wu, Junsong Yuan, Cem Keskin, Robert Wang*
24. DANBO: Disentangled Articulated Neural Body Representations via Graph Neural Networks, *Shih-Yang Su, Timur Bagautdinov, Helge Rhodin*
25. Three Things Everyone Should Know about Vision Transformers, *Hugo Touvron, Matthieu Cord, Alaaeldin El-Nouby, Jakob Verbeek, Hervé Jégou*
26. Telepresence Video Quality Assessment, *Zhenqiang Ying, Deepti Ghadiyaram, Alan Bovik*
27. OCR-Free Document Understanding Transformer, *Geewook Kim, Teakgyu Hong, Moonbin Yim, JeongYeon Nam, Jinyoung Park, Jinyeong Yim, Wonseok Hwang, Sangdoo Yun, Dongyoon Han, Seunghyun Park*
28. BlobGAN: Spatially Disentangled Scene Representations, *Dave Epstein, Taesung Park, Richard Zhang, Eli Shechtman, Alexei A. Efros*
29. Improving Robustness by Enhancing Weak Subnets, *Yong Guo, David Stutz, Bernt Schiele*
30. VecGAN: Image-to-Image Translation with Interpretable Latent Directions, *Yusuf Dalva, Said Fahri Altındış, Aysegül Dundar*
31. Rethinking Few-Shot Object Detection on a Multi-Domain Benchmark, *Kibok Lee, Hao Yang, Satyaki Chakraborty, Zhaowei Cai, Gurumurthy Swaminathan, Avinash Ravichandran, Onkar Dabeer*
32. Conditional-Flow NeRF: Accurate 3D Modelling with Reliable Uncertainty Quantification, *Jianxiong Shen, Antonio Agudo, Francesc Moreno-Noguer, Adria Ruiz*
33. Learning Pedestrian Group Representations for Multi-modal Trajectory Prediction, *Inhwan Bae, Jin-Hwi Park, Hae-Gon Jeon*
34. SHAPO: Implicit Representations for Multi-Object Shape, Appearance, and Pose Optimization, *Muhammad Zubair Irshad, Sergey Zakharov, Rareş Ambruş, Thomas Kollar, Zsolt Kira, Adrien Gaidon*
35. MoDA: Map Style Transfer for Self-Supervised Domain Adaptation of Embodied Agents, *Eun Sun Lee, Junho Kim, SangWon Park, Young Min Kim*
36. BASQ: Branch-Wise Activation-Clipping Search Quantization for Sub-4-Bit Neural Networks, *Han-Byul Kim, Eunhyeok Park, Sungjoo Yoo*
37. Multi-Person 3D Pose and Shape Estimation via Inverse Kinematics and Refinement, *Junuk Cha, Muhammad Saqlain, GeonU Kim, Mingyu Shin, Seungryul Baek*
38. Panoptic Scene Graph Generation, *Jingkang Yang, Yi Zhe Ang, Zujin Guo, Kaiyang Zhou, Wayne Zhang, Ziwei Liu*
39. Social-SSL: Self-Supervised Cross-Sequence Representation Learning Based on Transformers for Multi-agent Trajectory Prediction, *Li-Wu Tsao, Yan-Kai Wang, Hao-Siang Lin, Hong-Han Shuai, Lai-Kuan Wong, Wen-Huang Cheng*
40. Image-Based CLIP-Guided Essence Transfer, *Hila Chefer, Sagie Benaïm, Roni Paiss, Lior Wolf*
41. ECCV Caption: Correcting False Negatives by Collecting Machine-and-Human-Verified Image-Caption Associations for MS-COCO, *Sanghyuk Chun, Wonjae Kim, Song Park, Minsuk Chang, Seong Joon Oh*
42. GIPSO: Geometrically Informed Propagation for Online Adaptation in 3D LiDAR Segmentation, *Cristiano Saltori, Evgeny Krivosheev, Stéphane Lathuilière, Nicu Sebe, Fabio Galasso, Giuseppe Fiameni, Elisa Ricci, Fabio Poiesi*
43. Fast Two-View Motion Segmentation Using Christoffel Polynomials, *Bengisu Ozbay, Octavia Camps, Mario Sznajder*
44. Dual Perspective Network for Audio-Visual Event Localization, *Varshanth Rao, Md Ibrahim Khalil, Haoda Li, Peng Dai, Juwei Lu*
45. TREND: Truncated Generalized Normal Density Estimation of Inception Embeddings for GAN Evaluation, *Junghyuk Lee, Jong-Seok Lee*
46. Grounding Visual Representations with Texts for Domain Generalization, *Seonwoo Min, Nokyoung Park, Siwon Kim, Seunghyun Park, Jinkyu Kim*
47. FrequencyLowCut Pooling – Plug & Play against Catastrophic Overfitting, *Julia Grabinski, Steffen Jung, Janis Keuper, Margret Keuper*
48. GigaDepth: Learning Depth from Structured Light with Branching Neural Networks, *Simon Schreiberhuber, Jean-Baptiste Weibel, Timothy Patten, Markus Vincze*
49. Diverse Human Motion Prediction Guided by Multi-level Spatial-Temporal Anchors, *Sirui Xu, Yu-Xiong Wang, Liang-Yan Gui*
50. Hierarchical Average Precision Training for Pertinent Image Retrieval, *Elias Ramzi, Nicolas Audebert, Nicolas Thome, Clément Rambour, Xavier Bitot*
51. IntereStyle: Encoding an Interest Region for Robust StyleGAN Inversion, *Seung-Jun Moon, Gyeong-Moon Park*
52. Semi-Supervised Learning of Optical Flow by Flow Supervisor, *Woobin Im, Sebin Lee, Sung-Eui Yoon*
53. SPSN: Superpixel Prototype Sampling Network for RGB-D Salient Object Detection, *Minhyeok Lee, Chaewon Park, Suhwan Cho, Sangyoun Lee*
54. CANF-VC: Conditional Augmented Normalizing Flows for Video Compression, *Yung-Han Ho, Chih-Peng Chang, Peng-Yu Chen, Alessandro Gnutti, Wen-Hsiao Peng*
55. Context-Enhanced Stereo Transformer, *Wei Yu Guo, Zhaoshuo Li, Yongkui Yang, Zheng Wang, Russell H. Taylor, Mathias Unberath, Alan Yuille, Yingwei Li*
56. AlignSDF: Pose-Aligned Signed Distance Fields for Hand-Object Reconstruction, *Zerui Chen, Yana Hasson, Cordelia Schmid, Ivan Laptev*

57. Geometric Features Informed Multi-Person Human-Object Interaction Recognition in Videos, *Tanqiu Qiao, Qianhui Men, Frederick W. B. Li, Yoshiki Kubotani, Shigeo Morishima, Hubert P. H. Shum*
58. Streamable Neural Fields, *Junwoo Cho, Seungtae Nam, Daniel Rho, Jong Hwan Ko, Eunbyung Park*
59. Vote from the Center: 6 DoF Pose Estimation in RGB-D Images by Radial Keypoint Voting, *Yangzheng Wu, Mohsen Zand, Ali Etemad, Michael Greenspan*
60. PIP: Physical Interaction Prediction via Mental Simulation with Span Selection, *Jiafei Duan, Samson Yu, Soujanya Poria, Bihan Wen, Cheston Tan*
61. DeepMend: Learning Occupancy Functions to Represent Shape for Repair, *Nikolas Lamb, Sean Banerjee, Natasha Kholgade Banerjee*
62. DLCFT: Deep Linear Continual Fine-Tuning for General Incremental Learning, *Hyounguk Shon, Janghyeon Lee, Seung Hwan Kim, Junmo Kim*
63. Self-Supervised Classification Network, *Elad Amrani, Leonid Karlinsky, Alex Bronstein*
64. Towards Metrical Reconstruction of Human Faces, *Wojciech Zielonka, Timo Bolkart, Justus Thies*
65. Learned Vertex Descent: A New Direction for 3D Human Model Fitting, *Enric Corona, Gerard Pons-Moll, Guillem Alenyà, Francesc Moreno-Noguer*
66. Robust Landmark-Based Stent Tracking in X-Ray Fluoroscopy, *Luojie Huang, Yikang Liu, Li Chen, Eric Z. Chen, Xiao Chen, Shanhui Sun*
67. The Missing Link: Finding Label Relations across Datasets, *Jasper Uijlings, Thomas Mensink, Vittorio Ferrari*
68. Deep Ensemble Learning by Diverse Knowledge Distillation for Fine-Grained Object Classification, *Naoki Okamoto, Tsubasa Hiraoka, Takayoshi Yamashita, Hironobu Fujiyoshi*
69. Multimodal Object Detection via Probabilistic Ensembling, *Yi-Ting Chen, Jinghao Shi, Zelin Ye, Christoph Mertz, Deva Ramanan, Shu Kong*
70. Latent Space Smoothing for Individually Fair Representations, *Momchil Peychev, Anian Ruoss, Mislav Balunović, Maximilian Baader, Martin Vechev*
71. SUPR: A Sparse Unified Part-Based Human Representation, *Ahmed A. A. Osman, Timo Bolkart, Dimitrios Tzionas, Michael J. Black*
72. PolarMOT: How Far Can Geometric Relations Take Us in 3D Multi-Object Tracking? *Aleksandr Kim, Guillem Brasó, Aljoša Ošep, Laura Leal-Taixé*
73. Learning Instance-Specific Adaptation for Cross-Domain Segmentation, *Yuliang Zou, Zizhao Zhang, Chun-Liang Li, Han Zhang, Tomas Pfister, Jia-Bin Huang*
74. Motion Inspired Unsupervised Perception and Prediction in Autonomous Driving, *Mahyar Najibi, Jingwei Ji, Yin Zhou, Charles R. Qi, Xinchun Yan, Scott Ettinger, Dragomir Anguelov*
75. Graph Neural Network for Cell Tracking in Microscopy Videos, *Tal Ben-Haim, Tammy Riklin Raviv*
76. Correspondence Reweighted Translation Averaging, *Lalit Manam, Venu Madhav Govindu*
77. MotionCLIP: Exposing Human Motion Generation to CLIP Space, *Guy Tevet, Brian Gordon, Amir Hertz, Amit H. Bermano, Daniel Cohen-Or*
78. Learning Audio-Video Modalities from Image Captions, *Arsha Nagrani, Paul Hongsuck Seo, Bryan Seybold, Anja Hauth, Santiago Manen, Chen Sun, Cordelia Schmid*
79. Long-Tail Detection with Effective Class-Margins, *Jang Hyun Cho, Philipp Krähenbühl*
80. Unsupervised Domain Adaptation for One-Stage Object Detector Using Offsets to Bounding Box, *Jayeon Yoo, Inseop Chung, Nojun Kwak*
81. NeRF for Outdoor Scene Relighting, *Viktor Rudnev, Mohamed Elgharib, William Smith, Lingjie Liu, Vladislav Golyanik, Christian Theobalt*
82. A Comparative Study of Graph Matching Algorithms in Computer Vision, *Stefan Haller, Lorenz Feineis, Lisa Hutschenreiter, Florian Bernard, Carsten Rother, Dagmar Kainmüller, Paul Swoboda, Bogdan Savchynskyy*
83. Quantized GAN for Complex Music Generation from Dance Videos, *Ye Zhu, Kyle Olszewski, Yu Wu, Panos Achlioptas, Menglei Chai, Yan Yan, Sergey Tulyakov*
84. Visual Prompt Tuning, *Menglin Jia, Luming Tang, Bor-Chun Chen, Claire Cardie, Serge Belongie, Bharath Hariharan, Ser-Nam Lim*
85. Embedding Contrastive Unsupervised Features to Cluster in- and Out-of-Distribution Noise in Corrupted Image Datasets, *Paul Albert, Eric Arazo, Noel E. O'Connor, Kevin McGuinness*
86. Cross-Domain Ensemble Distillation for Domain Generalization, *Kyungmoon Lee, Sungyeon Kim, Suha Kwak*
87. Bridging the Visual Semantic Gap in VLN via Semantically Richer Instructions, *Joaquín Ossandón, Benjamín Earle, Alvaro Soto*
88. GAN Cocktail: Mixing GANs without Dataset Access, *Omri Avrahami, Dani Lischinski, Ohad Fried*
89. PREF: Predictability Regularized Neural Motion Fields, *Liangchen Song, Xuan Gong, Benjamin Planche, Meng Zheng, David Doermann, Junsong Yuan, Terrence Chen, Ziyang Wu*
90. CHORE: Contact, Human and Object REconstruction from a Single RGB Image, *Xianghui Xie, Bharat Lal Bhatnagar, Gerard Pons-Moll*
91. Controllable Video Generation through Global and Local Motion Dynamics, *Aram Davtyan, Paolo Favaro*
92. Online Segmentation of LiDAR Sequences: Dataset and Algorithm, *Romain Loiseau, Mathieu Aubry, Loïc Landrieu*
93. Inverted Pyramid Multi-Task Transformer for Dense Scene Understanding, *Hanrong Ye, Dan Xu*
94. Domain Knowledge-Informed Self-Supervised Representations for Workout Form Assessment, *Paritosh Parmar, Amol Gharat, Helge Rhodin*
95. Active Audio-Visual Separation of Dynamic Sound Sources, *Sagnik Majumder, Kristen Grauman*
96. NeuMan: Neural Human Radiance Field from a Single Video, *Wei Jiang, Kwang Moo Yi, Golnoosh Samei, Oncel Tuzel, Anurag Ranjan*
97. A Fast Knowledge Distillation Framework for Visual Recognition, *Zhiqiang Shen, Eric Xing*
98. Contrasting Quadratic Assignments for Set-Based Representation Learning, *Artem Moskalev, Ivan Sosnovik, Volker Fischer, Arnold Smeulders*
99. A Perturbation-Constrained Adversarial Attack for Evaluating the Robustness of Optical Flow, *Jenny Schmalfuss, Philipp Scholze, Andrés Bruhn*
100. Generative Negative Text Replay for Continual Vision-Language Pretraining, *Shipeng Yan, Lanqing Hong, Hang Xu, Jianhua Han, Tinne Tuytelaars, Zhenguo Li, Xuming He*
101. Privacy-Preserving Action Recognition via Motion Difference Quantization, *Sudhakar Kumawat, Hajime Nagahara*

-
- This image shows a full page of blank graph paper. The grid consists of small squares formed by thin black lines. There are 20 columns and 20 rows of squares. A thicker horizontal line runs across the middle of the page, dividing it into two equal halves of 10 rows each. This line serves as a baseline for writing. The rest of the page is filled with the standard grid pattern.

[illegible][illegible]

1330–1430 Poster Switch/Setup (Hall B)

1330–1530 Oral 1.B.1: Architecture, Training, and Optimization (Hall D; overflow in Halls K-M)

Papers in this session are in Poster Session 1.B

Format (9 min. presentation; 3 min. questions)

- [1330] Unpaired Image Translation via Vector Symbolic Architectures, *Justin Theiss, Jay Leverett, Daeil Kim, Aayush Prakash*
- [1342] Adaptive Token Sampling for Efficient Vision Transformers, *Mohsen Fayyaz, Soroush Abbasi Koohpayegani, Farnoush Rezaei Jafari, Sunando Sengupta, Hamid Reza Vaezi Joze, Eric Sommerlade, Hamed Pirsiavash, Jürgen Gall*
- [1354] Cross-Modal Knowledge Transfer without Task-Relevant Source Data, *Sk Miraj Ahmed, Suhas Lohit, Kuan-Chuan Peng, Michael J. Jones, Amit K. Roy-Chowdhury*
- [1406] The Challenges of Continuous Self-Supervised Learning, *Senthil Purushwalkam, Pedro Morgado, Abhinav Gupta*
- [1418] PACTran: PAC-Bayesian Metrics for Estimating the Transferability of Pretrained Models to Classification Tasks, *Nan Ding, Xi Chen, Tomer Levinboim, Soravit Changpinyo, Radu Soricut*
- [1430] Semidefinite Relaxations of Truncated Least-Squares in Robust Rotation Search: Tight or Not, *Liangzu Peng, Mahyar Fazlyab, René Vidal*
- [1442] Exploring Lottery Ticket Hypothesis in Spiking Neural Networks, *Youngeun Kim, Yuhang Li, Hyoungseob Park, Yeshwanth Venkatesha, Ruokai Yin, Priyadarshini Panda*
- [1454] Cartoon Explanations of Image Classifiers, *Stefan Kolek, Duc Anh Nguyen, Ron Levie, Joan Bruna, Gitta Kutyniok*

1330–1530 Oral 1.B.2: Shape From-X and Applications (Hall E; overflow in Halls F-H)

Papers in this session are in Poster Session 1.B

Format (9 min. presentation; 3 min. questions)

- [1330] Revisiting a kNN-based Image Classification System with High-capacity Storage, *Kengo Nakata, Youyang Ng, Daisuke Miyashita, Asuka Maki, Yu-Chieh Lin, Jun Deguchi*
- [1342] Physically-Based Editing of Indoor Scene Lighting from a Single Image, *Zhengqin Li, Jia Shi, Sai Bi, Rui Zhu, Kalyan Sunkavalli, Miloš Hašan, Zexiang Xu, Ravi Ramamoorthi, Manmohan Chandraker*
- [1354] A Level Set Theory for Neural Implicit Evolution under Explicit Flows, *Ishit Mehta, Manmohan Chandraker, Ravi Ramamoorthi*
- [1406] Implicit Field Supervision for Robust Non-rigid Shape Matching, *Ramana Sundararaman, Gautam Pai, Maks Ovsjanikov*
- [1418] Shape-Pose Disentanglement Using SE(3)-Equivariant Vector Neurons, *Oren Katzir, Dani Lischinski, Daniel Cohen-Or*
- [1430] Unsupervised Pose-Aware Part Decomposition for Man-Made Articulated Objects, *Yuki Kawana, Yusuke Mukuta, Tatsuya Harada*

1530–1630 Coffee & Refreshments (Lobby and Hall C)

1530–1730 Poster 1.B (Hall B)

- 1. Combating Label Distribution Shift for Active Domain Adaptation, *Sehyun Hwang, Sohyun Lee, Sungyeon Kim, Jungseul Ok, Suha Kwak*
- 2. Skeleton-Free Pose Transfer for Stylized 3D Characters, *Zhouyingcheng Liao, Jimei Yang, Jun Saito, Gerard Pons-Moll, Yang Zhou*
- 3. ConMatch: Semi-Supervised Learning with Confidence-Guided Consistency Regularization, *Jiwon Kim, Youngjo Min, Daehwan Kim, Gyuseong Lee, Junyoung Seo, Kwangrok Ryoo, Seungryong Kim*
- 4. Overcoming Shortcut Learning in a Target Domain by Generalizing Basic Visual Factors from a Source Domain, *Piyapat Saranrittichai, Chaithanya Kumar Mummadi, Claudia Blaiotta, Mauricio Munoz, Volker Fischer*
- 5. Unpaired Image Translation via Vector Symbolic Architectures, *Justin Theiss, Jay Leverett, Daeil Kim, Aayush Prakash*
- 6. Bi-directional Contrastive Learning for Domain Adaptive Semantic Segmentation, *Geon Lee, Chanho Eom, Wonkyung Lee, Hyekang Park, Bumsu Ham*
- 7. SelectionConv: Convolutional Neural Networks for Non-Rectilinear Image Data, *David Hart, Michael Whitney, Bryan Morse*
- 8. Static and Dynamic Concepts for Self-Supervised Video Representation Learning, *Rui Qian, Shuangrui Ding, Xian Liu, Dahua Lin*
- 9. SPot-the-Difference Self-Supervised Pre-training for Anomaly Detection and Segmentation, *Yang Zou, Jongheon Jeong, Latha Pemula, Dongqing Zhang, Onkar Dabeer*
- 10. Adaptive Token Sampling for Efficient Vision Transformers, *Mohsen Fayyaz, Soroush Abbasi Koohpayegani, Farnoush Rezaei Jafari, Sunando Sengupta, Hamid Reza Vaezi Joze, Eric Sommerlade, Hamed Pirsiavash, Jürgen Gall*
- 11. My View Is the Best View: Procedure Learning from Egocentric Videos, *Siddhant Bansal, Chetan Arora, C. V. Jawahar*
- 12. TL;DW? Summarizing Instructional Videos with Task Relevance & Cross-Modal Saliency, *Medhini Narasimhan, Arsha Nagrani, Chen Sun, Michael Rubinstein, Trevor Darrell, Anna Rohrbach, Cordelia Schmid*
- 13. Learning Hierarchy Aware Features for Reducing Mistake Severity, *Ashima Garg, Depanshu Sani, Saket Anand*
- 14. On the Angular Update and Hyperparameter Tuning of a Scale-Invariant Network, *Juseung Yun, Janghyeon Lee, Hyounguk Shon, Eojindl Yi, Seung Hwan Kim, Junmo Kim*
- 15. Cross-Modal Knowledge Transfer without Task-Relevant Source Data, *Sk Miraj Ahmed, Suhas Lohit, Kuan-Chuan Peng, Michael J. Jones, Amit K. Roy-Chowdhury*
- 16. ClearPose: Large-Scale Transparent Object Dataset and Benchmark, *Xiaotong Chen, Huijie Zhang, Zeren Yu, Anthony Opiari, Odest Chadwicke Jenkins*
- 17. AvatarPoser: Articulated Full-Body Pose Tracking from Sparse Motion Sensing, *Jiaxi Jiang, Paul Strel, Huajian Qiu, Andreas Fender, Larissa Laich, Patrick Snape, Christian Holz*
- 18. Custom Structure Preservation in Face Aging, *Guillermo Gomez-Trenado, Stéphane Lathuilière, Pablo Mesejo, Óscar Córdón*
- 19. Learning Where to Look – Generative NAS Is Surprisingly Efficient, *Jovita Lukasik, Steffen Jung, Margret Keuper*
- 20. The Challenges of Continuous Self-Supervised Learning, *Senthil Purushwalkam, Pedro Morgado, Abhinav Gupta*

21. AirDet: Few-Shot Detection without Fine-Tuning for Autonomous Exploration, *Bowen Li, Chen Wang, Pranay Reddy, Seungchan Kim, Sebastian Scherer*
22. Bridging the Domain Gap towards Generalization in Automatic Colorization, *Hyejin Lee, Daehee Kim, Daeun Lee, Jinkyu Kim, Jaekoo Lee*
23. Is Appearance Free Action Recognition Possible? *Filip Illic, Thomas Pock, Richard P. Wildes*
24. Object Detection As Probabilistic Set Prediction, *Georg Hess, Christoffer Petersson, Lennart Svensson*
25. DProST: Dynamic Projective Spatial Transformer Network for 6D Pose Estimation, *Jaewoo Park, Nam Ik Cho*
26. RC-MVSNet: Unsupervised Multi-View Stereo with Neural Rendering, *Di Chang, Aljaž Božič, Tong Zhang, Qingsong Yan, Yingcong Chen, Sabine Süsstrunk, Matthias Nießner*
27. Few-Shot Image Generation with Mixup-Based Distance Learning, *Chaerin Kong, Jeessoo Kim, Donghoon Han, Nojun Kwak*
28. PlaneFormers: From Sparse View Planes to 3D Reconstruction, *Samir Agarwala, Linyi Jin, Chris Rockwell, David F. Fouhey*
29. PoseGPT: Quantization-Based 3D Human Motion Generation and Forecasting, *Thomas Lucas, Fabien Baradel, Philippe Weinzaepfel, Grégory Rogez*
30. PACTran: PAC-Bayesian Metrics for Estimating the Transferability of Pretrained Models to Classification Tasks, *Nan Ding, Xi Chen, Tomer Levinboim, Soravit Changpinyo, Radu Soricut*
31. ARF: Artistic Radiance Fields, *Kai Zhang, Nick Kolkin, Sai Bi, Fujun Luan, Zexiang Xu, Eli Shechtman, Noah Snavely*
32. On Label Granularity and Object Localization, *Elijah Cole, Kimberly Wilber, Grant Van Horn, Xuan Yang, Marco Fornoni, Pietro Perona, Serge Belongie, Andrew Howard, Oisín Mac Aodha*
33. The Surprisingly Straightforward Scene Text Removal Method with Gated Attention and Region of Interest Generation: A Comprehensive Prominent Model Analysis, *Hyeonsu Lee, Chankyu Choi*
34. Order Learning Using Partially Ordered Data via Chainization, *Seon-Ho Lee, Chang-Su Kim*
35. LaLaLoc++: Global Floor Plan Comprehension for Layout Localisation in Unvisited Environments, *Henry Howard-Jenkins, Victor Adrian Prisacariu*
36. A Dataset Generation Framework for Evaluating Megapixel Image Classifiers & Their Explanations, *Gautam Machiraju, Sylvia Plevritis, Parag Mallick*
37. AdaNeRF: Adaptive Sampling for Real-Time Rendering of Neural Radiance Fields, *Andreas Kurz, Thomas Neff, Zhaoyang Lv, Michael Zollhöfer, Markus Steinberger*
38. Frozen CLIP Models Are Efficient Video Learners, *Ziyi Lin, Shijie Geng, Renrui Zhang, Peng Gao, Gerard de Melo, Xiaogang Wang, Jifeng Dai, Yu Qiao, Hongsheng Li*
39. R2L: Distilling Neural Radiance Field to Neural Light Field for Efficient Novel View Synthesis, *Huan Wang, Jian Ren, Zeng Huang, Kyle Olszewski, Menglei Chai, Yun Fu, Sergey Tulyakov*
40. Semidefinite Relaxations of Truncated Least-Squares in Robust Rotation Search: Tight or Not, *Liangzu Peng, Mahyar Fazlyab, René Vidal*
41. DVS-Voltmeter: Stochastic Process-Based Event Simulator for Dynamic Vision Sensors, *Songnan Lin, Ye Ma, Zhenhua Guo, Bihan Wen*
42. Zero-Shot Category-Level Object Pose Estimation, *Walter Goodwin, Sagar Vaze, Ioannis Havoutis, Ingmar Posner*
43. Studying Bias in GANs through the Lens of Race, *Vongani H. Maluleke, Neerja Thakkar, Tim Brooks, Ethan Weber, Trevor Darrell, Alexei A. Efros, Angjoo Kanazawa, Devin Guillory*
44. Structural Causal 3D Reconstruction, *Weiyang Liu, Zhen Liu, Liam Paull, Adrian Weller, Bernhard Schölkopf*
45. Prediction-Guided Distillation for Dense Object Detection, *Chenhongyi Yang, Mateusz Ochal, Amos Storkey, Elliot J. Crowley*
46. CLASTER: Clustering with Reinforcement Learning for Zero-Shot Action Recognition, *Shreyank N Gowda, Laura Sevilla-Lara, Frank Keller, Marcus Rohrbach*
47. Transform Your Smartphone into a DSLR Camera: Learning the ISP in the Wild, *Ardendu Shekhar Tripathi, Martin Danelljan, Samarth Shukla, Radu Timofte, Luc Van Gool*
48. Selective Query-Guided Debiasing for Video Corpus Moment Retrieval, *Sunjae Yoon, Ji Woo Hong, Eunseop Yoon, Dahyun Kim, Junyeong Kim, Hee Suk Yoon, Chang D. Yoo*
49. Learning Self-Prior for Mesh Denoising Using Dual Graph Convolutional Networks, *Shota Hattori, Tatsuya Yatagawa, Yutaka Ohtake, Hiromasa Suzuki*
50. Exploring Lottery Ticket Hypothesis in Spiking Neural Networks, *Youngeun Kim, Yuhang Li, Hyoungeob Park, Yeshwanth Venkatesha, Ruokai Yin, Priyadarshini Panda*
51. SpatialDETR: Robust Scalable Transformer-Based 3D Object Detection from Multi-View Camera Images with Global Cross-Sensor Attention, *Simon Doll, Richard Schulz, Lukas Schneider, Viviane Benzin, Markus Enzweiler, Hendrik P.A. Lensch*
52. TOCH: Spatio-Temporal Object-to-Hand Correspondence for Motion Refinement, *Keyang Zhou, Bharat Lal Bhatnagar, Jan Eric Lenssen, Gerard Pons-Moll*
53. Monitored Distillation for Positive Congruent Depth Completion, *Tian Yu Liu, Parth Agrawal, Allison Chen, Byung-Woo Hong, Alex Wong*
54. RelPose: Predicting Probabilistic Relative Rotation for Single Objects in the Wild, *Jason Y. Zhang, Deva Ramanan, Shubham Tulsiani*
55. Controllable and Guided Face Synthesis for Unconstrained Face Recognition, *Feng Liu, Minchul Kim, Anil Jain, Xiaoming Liu*
56. Learning Deep Non-Blind Image Deconvolution without Ground Truths, *Yuhui Quan, Zhuojie Chen, Huan Zheng, Hui Ji*
57. Sound-Guided Semantic Video Generation, *Seung Hyun Lee, Gyeongrok Oh, Wonmin Byeon, Chanyoung Kim, Won Jeong Ryoo, Sang Ho Yoon, Hyunjun Cho, Jihyun Bae, Jinkyu Kim, Sangpil Kim*
58. An Impartial Take to the CNN vs Transformer Robustness Contest, *Francesco Pinto, Philip H. S. Torr, Puneet K. Dokania*
59. Towards Sequence-Level Training for Visual Tracking, *Minji Kim, Seungkwan Lee, Jungseul Ok, Bohyung Han, Minsu Cho*
60. Cartoon Explanations of Image Classifiers, *Stefan Kolek, Duc Anh Nguyen, Ron Levie, Joan Bruna, Gitta Kutyniok*
61. Meta Spatio-Temporal Debiasing for Video Scene Graph Generation, *Li Xu, Haoxuan Qu, Jason Kuen, Jiuxiang Gu, Jun Liu*
62. Unleashing Transformers: Parallel Token Prediction with Discrete Absorbing Diffusion for Fast High-Resolution Image Generation from Vector-Quantized Codes, *Sam Bond-Taylor, Peter Hessey, Hiroshi Sasaki, Toby P. Breckon, Chris G. Willcocks*

63. GCISG: Guided Causal Invariant Learning for Improved Syn-to-Real Generalization, *Gilhyun Nam, Gyeongjae Choi, Kyungmin Lee*
64. Disentangling Architecture and Training for Optical Flow, *Deqing Sun, Charles Herrmann, Fitsum Reda, Michael Rubinstein, David J. Fleet, William T. Freeman*
65. Primitive-Based Shape Abstraction via Nonparametric Bayesian Inference, *Yuwei Wu, Weixiao Liu, Sipu Ruan, Gregory S. Chirikjian*
66. DLME: Deep Local-Flatness Manifold Embedding, *Zelin Zang, Siyuan Li, Di Wu, Ge Wang, Kai Wang, Lei Shang, Baigui Sun, Hao Li, Stan Z. Li*
67. Temporally Consistent Semantic Video Editing, *Yiran Xu, Badour AlBahar, Jia-Bin Huang*
68. Translating a Visual LEGO Manual to a Machine-Executable Plan, *Ruocheng Wang, Yunzhi Zhang, Jiayuan Mao, Chin-Yi Cheng, Jiajun Wu*
69. Q-FW: A Hybrid Classical-Quantum Frank-Wolfe for Quadratic Binary Optimization, *Alp Yurtsever, Tolga Birdal, Vladislav Golyanik*
70. Revisiting a kNN-based Image Classification System with High-capacity Storage, *Kengo Nakata, Youyang Ng, Daisuke Miyashita, Asuka Maki, Yu-Chieh Lin, Jun Deguchi*
71. CYBORGS: Contrastively Bootstrapping Object Representations by Grounding in Segmentation, *Renhao Wang, Hang Zhao, Yang Gao*
72. Joint Learning of Localized Representations from Medical Images and Reports, *Philip Müller, Georgios Kaissis, Congyu Zou, Daniel Rueckert*
73. AutoAvatar: Autoregressive Neural Fields for Dynamic Avatar Modeling, *Ziqian Bai, Timur Bagautdinov, Javier Romero, Michael Zollhöfer, Ping Tan, Shunsuke Saito*
74. Learning Semantic Segmentation from Multiple Datasets with Label Shifts, *Dongwan Kim, Yi-Hsuan Tsai, Yumin Suh, Masoud Faraki, Sparsh Garg, Manmohan Chandraker, Bohyung Han*
75. Free-Viewpoint RGB-D Human Performance Capture and Rendering, *Phong Nguyen-Ha, Nikolaos Sarafianos, Christoph Lassner, Janne Heikkilä, Tony Tung*
76. Zero-Shot Learning for Reflection Removal of Single 360-Degree Image, *Byeong-Ju Han, Jae-Young Sim*
77. Tackling Background Distraction in Video Object Segmentation, *Suhwan Cho, Heansung Lee, Minhyeok Lee, Chaewon Park, Sungjun Jang, Minjung Kim, Sangyoun Lee*
78. Image2Point: 3D Point-Cloud Understanding with 2D Image Pretrained Models, *Chenfang Xu, Shijia Yang, Tomer Galanti, Bichen Wu, Xiangyu Yue, Bohan Zhai, Wei Zhan, Peter Vajda, Kurt Keutzer, Masayoshi Tomizuka*
79. Spectrum-Aware and Transferable Architecture Search for Hyperspectral Image Restoration, *Wei He, Quanming Yao, Naoto Yokoya, Tatsumi Uezato, Hongyan Zhang, Liangpei Zhang*
80. A Level Set Theory for Neural Implicit Evolution under Explicit Flows, *Ishit Mehta, Manmohan Chandraker, Ravi Ramamoorthi*
81. AnimeCeleb: Large-Scale Animation CelebHeads Dataset for Head Reenactment, *Kangyeol Kim, Sunghyun Park, Jaeseong Lee, Sunghyo Chung, Junsoo Lee, Jaegul Choo*
82. Spotting Temporally Precise, Fine-Grained Events in Video, *James Hong, Haotian Zhang, Michaël Gharbi, Matthew Fisher, Kayvon Fatahalian*
83. Improving Test-Time Adaptation via Shift-Agnostic Weight Regularization and Nearest Source Prototypes, *Sungha Choi, Seunghan Yang, Seokeon Choi, Sungrack Yun*
84. Leveraging Action Affinity and Continuity for Semi-Supervised Temporal Action Segmentation, *Guodong Ding, Angela Yao*
85. Pose Forecasting in Industrial Human-Robot Collaboration, *Alessio Sampieri, Guido Maria D'Amely di Melendugno, Andrea Avogaro, Federico Cunico, Francesco Setti, Geri Skenderi, Marco Cristani, Fabio Galasso*
86. Source-Free Video Domain Adaptation by Learning Temporal Consistency for Action Recognition, *Yuecong Xu, Jianfei Yang, Haozhi Cao, Keyu Wu, Min Wu, Zhenghua Chen*
87. Deforming Radiance Fields with Cages, *Tianhan Xu, Tatsuya Harada*
88. MonteBoxFinder: Detecting and Filtering Primitives to Fit a Noisy Point Cloud, *Michaël Ramamonjisoa, Sinisa Stekovic, Vincent Lepetit*
89. Continual 3D Convolutional Neural Networks for Real-Time Processing of Videos, *Lukas Hedegaard, Alexandros Iosifidis*
90. Implicit Field Supervision for Robust Non-rigid Shape Matching, *Ramana Sundararaman, Gautam Pai, Maks Ovsjanikov*
91. Multi-Exit Semantic Segmentation Networks, *Alexandros Kouris, Stylianos I. Venieris, Stefanos Laskaridis, Nicholas Lane*
92. TM2T: Stochastic and Tokenized Modeling for the Reciprocal Generation of 3D Human Motions and Texts, *Chuan Guo, Xinxin Zuo, Sen Wang, Li Cheng*
93. NewsStories: Illustrating Articles with Visual Summaries, *Reuben Tan, Bryan A. Plummer, Kate Saenko, JP Lewis, Avneesh Sud, Thomas Leung*
94. Long Video Generation with Time-Agnostic VQGAN and Time-Sensitive Transformer, *Songwei Ge, Thomas Hayes, Harry Yang, Xi Yin, Guan Pang, David Jacobs, Jia-Bin Huang, Devi Parikh*
95. Autoregressive 3D Shape Generation via Canonical Mapping, *An-Chieh Cheng, Xueting Li, Sifei Liu, Min Sun, Ming-Hsuan Yang*
96. Monocular 3D Object Reconstruction with GAN Inversion, *Junzhe Zhang, Daxuan Ren, Zhongang Cai, Chai Kiat Yeo, Bo Dai, Chen Change Loy*
97. HRDA: Context-Aware High-Resolution Domain-Adaptive Semantic Segmentation, *Lukas Hoyer, Dengxin Dai, Luc Van Gool*
98. TIPS: Text-Induced Pose Synthesis, *Prasun Roy, Subhankar Ghosh, Saumik Bhattacharya, Umapada Pal, Michael Blumenstein*
99. Decouple-and-Sample: Protecting Sensitive Information in Task Agnostic Data Release, *Abhishek Singh, Ethan Garza, Ayush Chopra, Praneeth Vepakomma, Vivek Sharma, Ramesh Raskar*
100. Shape-Pose Disentanglement Using SE(3)-Equivariant Vector Neurons, *Oren Katzir, Dani Lischinski, Daniel Cohen-Or*
101. Temporal and Cross-Modal Attention for Audio-Visual Zero-Shot Learning, *Otniel-Bogdan Mercea, Thomas Hummel, A. Sophia Koepke, Zeynep Akata*
102. Texturify: Generating Textures on 3D Shape Surfaces, *Yawar Siddiqui, Justus Thies, Fangchang Ma, Qi Shan, Matthias Nießner, Angela Dai*
103. Manifold Adversarial Learning for Cross-Domain 3D Shape Representation, *Hao Huang, Cheng Chen, Yi Fang*

-
- A full-page sheet of white graph paper with a light gray grid. The grid consists of small squares, approximately 10 units wide by 20 units high. There are no margins or additional markings on the page.

[illegible][illegible]

Wednesday, October 26

0700–1830 Registration (Pavillion 2 Lobby)

0800–0900 Coffee & Refreshments (Lobby and Hall C)

0800–0900 Poster Setup (Hall B)

0900–1630 Exhibits (Hall C)

- See online for exhibitor list and map.

0900–1530 Demos (Hall C Demo Area)

- Artificial Intelligence Methods for Historical Documents Transcription, *Claudio Ronchetti (ENEA, Italy)*
- Open World Interactive Object Segmentation in 3D Point Clouds, *Theodora Kontogianni, Francis Engelmann (ETH-Zurich, Switzerland)*
- Accurate and Efficient Absolute 3D Human Pose Estimation Trained on Dozens of Datasets, *Istvan Sarandi (RWTH Aachen Univ., Germany)*
- Demonstrating AvatarPoser for Full-Body Pose Tracking from Sparse Motion Sensing, *Jiaxi Jiang, Paul Streli, Huqian Qiu, Andreas Fender (ETH-Zurich, Switzerland), Larissa Laich, Patrick Snape (Meta, Switzerland), Christian Holz (ETH-Zurich, Switzerland)*

0900–1100 Oral 2.A.1: Image/Video Synthesis and Generative Models (Hall D; overflow in Halls K-M)

Papers in this session are in Poster Session 2.A

Format (9 min. presentation: 3 min. questions)

- [0900] Semantic-Aware Implicit Neural Audio-Driven Video Portrait Generation, *Xian Liu, Yinghao Xu, Qianyi Wu, Hang Zhou, Wayne Wu, Bolei Zhou*
- [0912] Make-a-Scene: Scene-Based Text-to-Image Generation with Human Priors, *Oran Gafni, Adam Polyak, Oran Ashual, Shelly Sheynin, Devi Parikh, Yaniv Taigman*
- [0924] RFNet-4D: Joint Object Reconstruction and Flow Estimation from 4D Point Clouds, *Tuan-Anh Vu, Thanh Nguyen, Binh-Son Hua, Quang-Hieu Pham, Sai-Kit Yeung*
- [0936] Text2LIVE: Text-Driven Layered Image and Video Editing, *Omer Bar-Tal, Dolev Ofri-Amar, Rafail Fridman, Yoni Kasten, Tali Dekel*
- [0948] Exploring Gradient-Based Multi-directional Controls in GANs, *Zikun Chen, Ruowei Jiang, Brendan Duke, Han Zhao, Parham Aarabi*
- [1000] 3D-Aware Indoor Scene Synthesis with Depth Priors, *Zifan Shi, Yujun Shen, Jiapeng Zhu, Dit-Yan Yeung, Qifeng Chen*
- [1012] Generative Multiplane Images: Making a 2D GAN 3D-Aware, *Xiaoming Zhao, Fangchang Ma, David Güera, Zhile Ren, Alexander G. Schwing, Alex Colburn*
- [1024] Layered Controllable Video Generation, *Jiahui Huang, Yuhe Jin, Kwang Moo Yi, Leonid Sigal*

0900–1100 Oral 2.A.2: Faces, Bodies, Gestures, and Pose (Hall E; overflow in Halls F-H)

Papers in this session are in Poster Session 2.A

Format (9 min. presentation: 3 min. questions)

- [0900] Pose-NDF: Modeling Human Pose Manifolds with Neural Distance Fields, *Garvita Tiwari, Dimitrije Antić, Jan Eric Lenssen, Nikolaos Sarafianos, Tony Tung, Gerard Pons-Moll*
- [0912] CLIFF: Carrying Location Information in Full Frames into Human Pose and Shape Estimation, *Zhihao Li, Jianzhuang Liu, Zhensong Zhang, Songcen Xu, Youliang Yan*
- [0924] Grasp'D: Differentiable Contact-Rich Grasp Synthesis for Multi-Fingered Hands, *Dylan Turpin, Liqun Wang, Eric Heiden, Yun-Chun Chen, Miles Macklin, Stavros Tsogkas, Sven Dickinson, Animesh Garg*
- [0936] PressureVision: Estimating Hand Pressure from a Single RGB Image, *Patrick Grady, Chengcheng Tang, Samarth Brahmabhatt, Christopher D. Twigg, Chengde Wan, James Hays, Charles C. Kemp*
- [0948] Pose for Everything: Towards Category-Agnostic Pose Estimation, *Lumin Xu, Sheng Jin, Wang Zeng, Wentao Liu, Chen Qian, Wanli Ouyang, Ping Luo, Xiaogang Wang*
- [1000] Multi-Domain Learning for Updating Face Anti-Spoofing Models, *Xiao Guo, Yaojie Liu, Anil Jain, Xiaoming Liu*

1100–1200 Coffee & Refreshments (Lobby and Hall C)

1100–1330 Mentoring Session (Halls K-M)

1100–1330 Industry Track (Hall E)

- See the ECCV program website or scan the QR code for the list of speakers and schedule.



1100–1330 Poster 2.A (Hall B)

1. MFIM: Megapixel Facial Identity Manipulation, *Sanghyeon Na*
2. VoViT: Low Latency Graph-Based Audio-Visual Voice Separation Transformer, *Juan F. Montesinos, Venkatesh S. Kadandale, Gloria Haro*
3. Tracking Every Thing in the Wild, *Siyuan Li, Martin Danelljan, Henghui Ding, Thomas E. Huang, Fisher Yu*
4. Almost-Orthogonal Layers for Efficient General-Purpose Lipschitz Networks, *Bernd Prach, Christoph H. Lampert*
5. Pose-NDF: Modeling Human Pose Manifolds with Neural Distance Fields, *Garvita Tiwari, Dimitrije Antić, Jan Eric Lenssen, Nikolaos Sarafianos, Tony Tung, Gerard Pons-Moll*
6. Bilateral Normal Integration, *Xu Cao, Hiroaki Santo, Boxin Shi, Fumio Okura, Yasuyuki Matsushita*
7. Constrained Mean Shift Using Distant Yet Related Neighbors for Representation Learning, *K L Navaneet, Soroush Abbasi Koohpayegani, Ajinkya Tejankar, Kossar Pourahmadi, Akshayvarun Subramanya, Hamed Pirsiavash*
8. Self-Supervised Sparse Representation for Video Anomaly Detection, *Jih-Ciang Wu, He-Yen Hsieh, Ding-Jie Chen, Chiou-Shann Fuh, Tyng-Luh Liu*
9. Dress Code: High-Resolution Multi-Category Virtual Try-On, *Davide Morelli, Matteo Fincato, Marcella Cornia, Federico Landi, Fabio Cesari, Rita Cucchiara*
10. CLIFF: Carrying Location Information in Full Frames into Human Pose and Shape Estimation, *Zhihao Li, Jianzhuang Liu, Zhensong Zhang, Songcen Xu, Youliang Yan*

11. STEEX: Steering Counterfactual Explanations with Semantics, *Paul Jacob, Éloi Zablocki, Hédi Ben-Younes, Mickaël Chen, Patrick Pérez, Matthieu Cord*
12. Graph-Constrained Contrastive Regularization for Semi-Weakly Volumetric Segmentation, *Simon Reiß, Constantin Seibold, Alexander Freytag, Erik Rodner, Rainer Stiefelhofen*
13. A Contrastive Objective for Learning Disentangled Representations, *Jonathan Kahana, Yedid Hoshen*
14. Less than Few: Self-Shot Video Instance Segmentation, *Pengwan Yang, Yuki M. Asano, Pascal Mettes, Cees G. M. Snoek*
15. Grasp'D: Differentiable Contact-Rich Grasp Synthesis for Multi-Fingered Hands, *Dylan Turpin, Liquan Wang, Eric Heiden, Yun-Chun Chen, Miles Macklin, Stavros Tsogkas, Sven Dickinson, Animesh Garg*
16. Vector Quantized Image-to-Image Translation, *Yu-Jie Chen, Shin-I Cheng, Wei-Chen Chiu, Hung-Yu Tseng, Hsin-Ying Lee*
17. Large Scale Real-World Multi-person Tracking, *Bing Shuai, Alessandro Bergamo, Uta Büchler, Andrew Berneshawi, Alyssa Boden, Joseph Tighe*
18. Supervised Attribute Information Removal and Reconstruction for Image Manipulation, *Nannan Li, Bryan A. Plummer*
19. Learning Local Implicit Fourier Representation for Image Warping, *Jaewon Lee, Kwang Pyo Choi, Kyong Hwan Jin*
20. PressureVision: Estimating Hand Pressure from a Single RGB Image, *Patrick Grady, Chengcheng Tang, Samarth Brahmabhatt, Christopher D. Twigg, Chengde Wan, James Hays, Charles C. Kemp*
21. Emotion-Aware Multi-View Contrastive Learning for Facial Emotion Recognition, *Daeha Kim, Byung Cheol Song*
22. FCAF3D: Fully Convolutional Anchor-Free 3D Object Detection, *Danila Rukhovich, Anna Vorontsova, Anton Konushin*
23. Any-Resolution Training for High-Resolution Image Synthesis, *Lucy Chai, Michaël Gharbi, Eli Shechtman, Phillip Isola, Richard Zhang*
24. QISTA-ImageNet: A Deep Compressive Image Sensing Framework Solving ℓ_q -Norm Optimization Problem, *Gang-Xuan Lin, Shih-Wei Hu, Chun-Shien Lu*
25. Pose for Everything: Towards Category-Agnostic Pose Estimation, *Lumin Xu, Sheng Jin, Wang Zeng, Wentao Liu, Chen Qian, Wanli Ouyang, Ping Luo, Xiaogang Wang*
26. Bitwidth-Adaptive Quantization-Aware Neural Network Training: A Meta-Learning Approach, *Jiseok Youn, Jaehun Song, Hyung-Sin Kim, Saewoong Bahk*
27. SWFormer: Sparse Window Transformer for 3D Object Detection in Point Clouds, *Pei Sun, Mingxing Tan, Weiyue Wang, Chenxi Liu, Fei Xia, Zhaoqi Leng, Dragomir Anguelov*
28. 2D GANs Meet Unsupervised Single-View 3D Reconstruction, *Feng Liu, Xiaoming Liu*
29. Towards Ultra Low Latency Spiking Neural Networks for Vision and Sequential Tasks Using Temporal Pruning, *Sayed Shafayet Chowdhury, Nitin Rath, Kaushik Roy*
30. Multi-Domain Learning for Updating Face Anti-Spoofing Models, *Xiao Guo, Yaojie Liu, Anil Jain, Xiaoming Liu*
31. Gradient-Based Uncertainty for Monocular Depth Estimation, *Julia Hornauer, Vasileios Belagiannis*
32. KeypointNeRF: Generalizing Image-Based Volumetric Avatars Using Relative Spatial Encoding of Keypoints, *Marko Mihajlovic, Aayush Bansal, Michael Zollhöfer, Siyu Tang, Shunsuke Saito*
33. Totems: Physical Objects for Verifying Visual Integrity, *Jingwei Ma, Lucy Chai, Minyoung Huh, Tongzhou Wang, Ser-Nam Lim, Phillip Isola, Antonio Torralba*
34. Realistic Blur Synthesis for Learning Image Deblurring, *Jaesung Rim, Geonung Kim, Jungeon Kim, Junyong Lee, Seungyong Lee, Sunghyun Cho*
35. Semantic-Aware Implicit Neural Audio-Driven Video Portrait Generation, *Xian Liu, Yinghao Xu, Qianyi Wu, Hang Zhou, Wayne Wu, Bolei Zhou*
36. Proposal-Free Temporal Action Detection via Global Segmentation Mask Learning, *Sauradip Nag, Xiatian Zhu, Yi-Zhe Song, Tao Xiang*
37. Speaker-Adaptive Lip Reading with User-Dependent Padding, *Minsu Kim, Hyunjun Kim, Yong Man Ro*
38. Editable Indoor Lighting Estimation, *Henrique Weber, Mathieu Garon, Jean-François Lalonde*
39. Cross-Modal Prototype Driven Network for Radiology Report Generation, *Jun Wang, Abhir Bhalerao, Yulan He*
40. Make-a-Scene: Scene-Based Text-to-Image Generation with Human Priors, *Oran Gafni, Adam Polyak, Oron Ashual, Shelly Sheynin, Devi Parikh, Yaniv Taigman*
41. FingerprintNet: Synthesized Fingerprints for Generated Image Detection, *Yonghyun Jeong, Doyeon Kim, Youngmin Ro, Pyounggeon Kim, Jongwon Choi*
42. Cost Aggregation with 4D Convolutional Swin Transformer for Few-Shot Segmentation, *Sunghwan Hong, Seokju Cho, Jisu Nam, Stephen Lin, Seungryong Kim*
43. Deep Hash Distillation for Image Retrieval, *Young Kyun Jang, Geonmo Gu, Byungsoo Ko, Isaac Kang, Nam Ik Cho*
44. Inpainting at Modern Camera Resolution by Guided PatchMatch with Auto-Curation, *Lingzhi Zhang, Connelly Barnes, Kevin Wampler, Sohrab Amirghodsi, Eli Shechtman, Zhe Lin, Jianbo Shi*
45. Style-Hallucinated Dual Consistency Learning for Domain Generalized Semantic Segmentation, *Yuyang Zhao, Zhun Zhong, Na Zhao, Nicu Sebe, Gim Hee Lee*
46. Hunting Group Clues with Transformers for Social Group Activity Recognition, *Masato Tamura, Rahul Vishwakarma, Ravigopal Vennelakanti*
47. Multi-Domain Multi-Definition Landmark Localization for Small Datasets, *David Ferman, Gaurav Bharaj*
48. Active Learning Strategies for Weakly-Supervised Object Detection, *Huy V. Vo, Oriane Siméoni, Spyros Gidaris, Andrei Bursuc, Patrick Pérez, Jean Ponce*
49. CoSMix: Compositional Semantic Mix for Domain Adaptation in 3D LiDAR Segmentation, *Cristiano Saltori, Fabio Galasso, Giuseppe Fiameni, Nicu Sebe, Elisa Ricci, Fabio Poiesi*
50. RFNet-4D: Joint Object Reconstruction and Flow Estimation from 4D Point Clouds, *Tuan-Anh Vu, Thanh Nguyen, Binh-Son Hua, Quang-Hieu Pham, Sai-Kit Yeung*
51. Optimizing Image Compression via Joint Learning with Denoising, *Ka Leong Cheng, Yueqi Xie, Qifeng Chen*
52. FLEX: Extrinsic Parameters-Free Multi-View 3D Human Motion Reconstruction, *Brian Gordon, Sigal Raab, Guy Azov, Raja Giryes, Daniel Cohen-Or*
53. Improving GANs for Long-Tailed Data through Group Spectral Regularization, *Harsh Rangwani, Naman Jaswani, Tejan Karmali, Varun Jampani, R. Venkatesh Babu*
54. SSBNet: Improving Visual Recognition Efficiency by Adaptive Sampling, *Ho Man Kwan, Shenghui Song*

55. Planes vs. Chairs: Category-Guided 3D Shape Learning without Any 3D Cues, *Zixuan Huang, Stefan Stojanov, Anh Thai, Varun Jampani, James M. Rehg*
56. X-DETR: A Versatile Architecture for Instance-Wise Vision-Language Tasks, *Zhaowei Cai, Gukyeong Kwon, Avinash Ravichandran, Erhan Bas, Zhuowen Tu, Rahul Bhotika, Stefano Soatto*
57. VTC: Improving Video-Text Retrieval with User Comments, *Laura Hanu, James Thewlis, Yuki M. Asano, Christian Rupprecht*
58. Weakly-Supervised Stitching Network for Real-World Panoramic Image Generation, *Dae-Young Song, Geonsoo Lee, HeeKyung Lee, Gi-Mun Um, Donghyeon Cho*
59. BRACE: The Breakdancing Competition Dataset for Dance Motion Synthesis, *Davide Moltisanti, Jinyi Wu, Bo Dai, Chen Change Loy*
60. Text2LIVE: Text-Driven Layered Image and Video Editing, *Omer Bar-Tal, Dolev Ofri-Amar, Rafail Fridman, Yoni Kasten, Tali Dekel*
61. Realistic One-Shot Mesh-Based Head Avatars, *Taras Khakhulin, Vanessa Sklyarova, Victor Lempitsky, Egor Zakharov*
62. MaCLR: Motion-Aware Contrastive Learning of Representations for Videos, *Fanyi Xiao, Joseph Tighe, Davide Modolo*
63. Semi-Supervised Temporal Action Detection with Proposal-Free Masking, *Sauradip Nag, Xiatian Zhu, Yi-Zhe Song, Tao Xiang*
64. FurryGAN: High Quality Foreground-Aware Image Synthesis, *Jeongmin Bae, Mingi Kwon, Youngjung Uh*
65. ERA: Enhanced Rational Activations, *Martin Trimmel, Mihai Zanfir, Richard Hartley, Cristian Sminchisescu*
66. Domain Adaptive Video Segmentation via Temporal Pseudo Supervision, *Yun Xing, Dayan Guan, Jiaying Huang, Shijian Lu*
67. Video Graph Transformer for Video Question Answering, *Junbin Xiao, Pan Zhou, Tat-Seng Chua, Shuicheng Yan*
68. The Anatomy of Video Editing: A Dataset and Benchmark Suite for AI-Assisted Video Editing, *Dawit Mureja Argaw, Fabian Caba, Joon-Young Lee, Markus Woodson, In So Kweon*
69. SCAM! Transferring Humans between Images with Semantic Cross Attention Modulation, *Nicolas Dufour, David Picard, Vicky Kalogeiton*
70. Exploring Gradient-Based Multi-directional Controls in GANs, *Zikun Chen, Ruowei Jiang, Brendan Duke, Han Zhao, Parham Aarabi*
71. Scene Text Recognition with Permuted Autoregressive Sequence Models, *Darwin Bautista, Rowel Atienza*
72. Robust Visual Tracking by Segmentation, *Mathieu Paul, Martin Danelljan, Christoph Mayer, Luc Van Gool*
73. A Sketch Is Worth a Thousand Words: Image Retrieval with Text and Sketch, *Patsorn Sangkloy, Wittawat Jitkrittum, Diyi Yang, James Hays*
74. Towards Efficient and Effective Self-Supervised Learning of Visual Representations, *Sravanti Addepalli, Kaushal Bhogale, Priyam Dey, R. Venkatesh Babu*
75. Learned Variational Video Color Propagation, *Markus Hofinger, Erich Kobler, Alexander Effland, Thomas Pock*
76. Dual-Domain Self-Supervised Learning and Model Adaption for Deep Compressive Imaging, *Yuhui Quan, Xinran Qin, Tongyao Pang, Hui Ji*
77. SLiDE: Self-Supervised LiDAR De-Snowing through Reconstruction Difficulty, *Gwangtak Bae, Byungjun Kim, Seongyong Ahn, Jihong Min, Inwook Shim*
78. Ensemble Knowledge Guided Sub-network Search and Fine-Tuning for Filter Pruning, *Seunghyun Lee, Byung Cheol Song*
79. Intrinsic Neural Fields: Learning Functions on Manifolds, *Lukas Koestler, Daniel Grittner, Michael Moeller, Daniel Cremers, Zorah Löhner*
80. 3D-Aware Indoor Scene Synthesis with Depth Priors, *Zifan Shi, Yujun Shen, Jiapeng Zhu, Dit-Yan Yeung, Qifeng Chen*
81. Masked Autoencoders for Point Cloud Self-Supervised Learning, *Yatian Pang, Wenxiao Wang, Francis E.H. Tay, Wei Liu, Yonghong Tian, Li Yuan*
82. How to Synthesize a Large-Scale and Trainable Micro-Expression Dataset? *Yuchi Liu, Zhongdao Wang, Tom Gedeon, Liang Zheng*
83. Learning Degradation Representations for Image Deblurring, *Dasong Li, Yi Zhang, Ka Chun Cheung, Xiaogang Wang, Hongwei Qin, Hongsheng Li*
84. Dynamic 3D Scene Analysis by Point Cloud Accumulation, *Shengyu Huang, Zan Gojcic, Jiahui Huang, Andreas Wieser, Konrad Schindler*
85. Share with Thy Neighbors: Single-View Reconstruction by Cross-Instance Consistency, *Tom Monnier, Matthew Fisher, Alexei A. Efros, Mathieu Aubry*
86. LWGNet – Learned Wirtinger Gradients for Fourier Ptychographic Phase Retrieval, *Atreyee Saha, Salman S. Khan, Sagar Sehrawat, Sanjana S. Prabhu, Shanti Bhattacharya, Kaushik Mitra*
87. 4DContrast: Contrastive Learning with Dynamic Correspondences for 3D Scene Understanding, *Yujin Chen, Matthias Nießner, Angela Dai*
88. RAWtoBit: A Fully End-to-End Camera ISP Network, *Wooseok Jeong, Seung-Won Jung*
89. When Deep Classifiers Agree: Analyzing Correlations between Learning Order and Image Statistics, *Iuliia Pliushch, Martin Mundt, Nicolas Lupp, Visvanathan Ramesh*
90. Generative Multiplane Images: Making a 2D GAN 3D-Aware, *Xiaoming Zhao, Fangchang Ma, David Güera, Zhile Ren, Alexander G. Schwing, Alex Colburn*
91. Object Discovery and Representation Networks, *Olivier J. Hénaff, Skanda Koppula, Evan Shelhamer, Daniel Zoran, Andrew Jaegle, Andrew Zisserman, João Carreira, Relja Arandjelović*
92. OIMNet++: Prototypical Normalization and Localization-Aware Learning for Person Search, *Sanghoon Lee, Youngmin Oh, Donghyeon Baek, Junghyup Lee, Bumsu Ham*
93. PS-NeRF: Neural Inverse Rendering for Multi-View Photometric Stereo, *Wenqi Yang, Guanying Chen, Chaofeng Chen, Zhenfang Chen, Kwan-Yee K. Wong*
94. EgoBody: Human Body Shape and Motion of Interacting People from Head-Mounted Devices, *Siwei Zhang, Qianli Ma, Yan Zhang, Zhiyin Qian, Taein Kwon, Marc Pollefeys, Federica Bogo, Siyu Tang*
95. Exploiting the Local Parabolic Landscapes of Adversarial Losses to Accelerate Black-Box Adversarial Attack, *Hoang Tran, Dan Lu, Guannan Zhang*
96. GradAuto: Energy-Oriented Attack on Dynamic Neural Networks, *Jianhong Pan, Qichen Zheng, Zhipeng Fan, Hossein Rahmani, Qihong Ke, Jun Liu*
97. Equivariant Hypergraph Neural Networks, *Jinwoo Kim, Saeyoon Oh, Sungjun Cho, Seunghoon Hong*

- [illegible]

[illegible][illegible][illegible]

1330–1430 Poster Switch/Setup (Hall B)

1330–1530 Oral 2.B.1: Scene, Action, and Video Understanding (Hall D; overflow in Halls K-M)

Papers in this session are in Poster Session 2.B

Format (9 min. presentation; 3 min. questions)

- [1330] ECLIPSE: Efficient Long-Range Video Retrieval Using Sight and Sound, *Yan-Bo Lin, Jie Lei, Mohit Bansal, Gedas Bertasius*
- [1342] GraphVid: It Only Takes a Few Nodes to Understand a Video, *Eitan Kosman, Dotan Di Castro*
- [1354] PrivHAR: Recognizing Human Actions from Privacy-Preserving Lens, *Carlos Hinojosa, Miguel Marquez, Henry Arguello, Ehsan Adeli, Li Fei-Fei, Juan Carlos Niebles*
- [1406] Flow Graph to Video Grounding for Weakly-Supervised Multi-step Localization, *Nikita Dvornik, Isma Hadji, Hai Pham, Dhaivat Bhatt, Brais Martinez, Afsaneh Fazly, Allan D. Jepson*
- [1418] Bi-PointFlowNet: Bidirectional Learning for Point Cloud Based Scene Flow Estimation, *Wencan Cheng, Jong Hwan Ko*
- [1430] Bayesian Tracking of Video Graphs Using Joint Kalman Smoothing and Registration, *Aditi Basu Bal, Ramy Mounir, Sathyanarayanan Aakur, Sudeep Sarkar, Anuj Srivastava*
- [1442] Uncertainty-DTW for Time Series and Sequences, *Lei Wang, Piotr Koniusz*

1330–1530 Oral 2.B.2: Low-Level Vision, Segmentation, and Synthesis (Hall E; overflow in Halls F-H)

Papers in this session are in Poster Session 2.B

Format (9 min. presentation; 3 min. questions)

- [1330] InfiniteNature-Zero: Learning Perpetual View Generation of Natural Scenes from Single Images, *Zhengqi Li, Qianqian Wang, Noah Snavely, Angjoo Kanazawa*
- [1342] Perceptual Artifacts Localization for Inpainting, *Lingzhi Zhang, Yuqian Zhou, Connelly Barnes, Sohrab Amirghodsi, Zhe Lin, Eli Shechtman, Jianbo Shi*
- [1354] Semantic-Aware Fine-Grained Correspondence, *Yingdong Hu, Renhao Wang, Kaifeng Zhang, Yang Gao*
- [1406] Secrets of Event-Based Optical Flow, *Shintaro Shiba, Yoshimitsu Aoki, Guillermo Gallego*
- [1418] Event-Based Fusion for Motion Deblurring with Cross-Modal Attention, *Lei Sun, Christos Sakaridis, Jingyun Liang, Qi Jiang, Kailun Yang, Peng Sun, Yaozu Ye, Kaiwei Wang, Luc Van Gool*

1530–1630 Coffee & Refreshments (Lobby and Hall C)

1530–1730 Industry Track (Halls K-M)

- See the ECCV program website or scan the QR code for the list of speakers and schedule.



1530–1730 Poster 2.B (Hall B)

- 1. One-Trimap Video Matting, *Hongje Seong, Seoung Wug Oh, Brian Price, Euntai Kim, Joon-Young Lee*
- 2. Tackling Long-Tailed Category Distribution under Domain Shifts, *Xiao Gu, Yao Guo, Zeju Li, Jianing Qiu, Qi Dou, Yuxuan Liu, Benny Lo, Guang-Zhong Yang*

- 3. Paint2Pix: Interactive Painting Based Progressive Image Synthesis and Editing, *Jaskirat Singh, Liang Zheng, Cameron Smith, Jose Echevarria*
- 4. Fusing Local Similarities for Retrieval-Based 3D Orientation Estimation of Unseen Objects, *Chen Zhao, Yinlin Hu, Mathieu Salzmann*
- 5. End-to-End Visual Editing with a Generatively Pre-trained Artist, *Andrew Brown, Cheng-Yang Fu, Omkar Parkhi, Tamara L. Berg, Andrea Vedaldi*
- 6. COUCH: Towards Controllable Human-Chair Interactions, *Xiaohan Zhang, Bharat Lal Bhatnagar, Sebastian Starke, Vladimir Guzun, Gerard Pons-Moll*
- 7. Concurrent Subsidiary Supervision for Unsupervised Source-Free Domain Adaptation, *Jogendra Nath Kundu, Suvaansh Bhambri, Akshay Kulkarni, Hiran Sarkar, Varun Jampani, R. Venkatesh Babu*
- 8. StretchBEV: Stretching Future Instance Prediction Spatially and Temporally, *Adil Kaan Akan, Fatma Güney*
- 9. Locally Varying Distance Transform for Unsupervised Visual Anomaly Detection, *Wen-Yan Lin, Zhonghang Liu, Siying Liu*
- 10. ECLIPSE: Efficient Long-Range Video Retrieval Using Sight and Sound, *Yan-Bo Lin, Jie Lei, Mohit Bansal, Gedas Bertasius*
- 11. BayesCap: Bayesian Identity Cap for Calibrated Uncertainty in Frozen Neural Networks, *Uddeshya Upadhyay, Shyamgopal Karthik, Yanbei Chen, Massimiliano Mancini, Zeynep Akata*
- 12. Learning Online Multi-Sensor Depth Fusion, *Erik Sandström, Martin R. Oswald, Suryansh Kumar, Silvan Weder, Fisher Yu, Cristian Sminchisescu, Luc Van Gool*
- 13. Target-Absent Human Attention, *Zhibo Yang, Sounak Mondal, Seoyoung Ahn, Gregory Zelinsky, Minh Hoai, Dimitris Samaras*
- 14. CPrune: Compiler-Informed Model Pruning for Efficient Target-Aware DNN Execution, *Taeho Kim, Yongin Kwon, Jemin Lee, Taeho Kim, Sangtae Ha*
- 15. Teaching with Soft Label Smoothing for Mitigating Noisy Labels in Facial Expressions, *Tohar Lukov, Na Zhao, Gim Hee Lee, Ser-Nam Lim*
- 16. Hierarchical Semantic Regularization of Latent Spaces in StyleGANs, *Tejan Karmali, Rishubh Parihar, Susmit Agrawal, Harsh Rangwani, Varun Jampani, Maneesh Singh, R. Venkatesh Babu*
- 17. Dense Gaussian Processes for Few-Shot Segmentation, *Joakim Johnander, Johan Edstedt, Michael Felsberg, Fahad Shahbaz Khan, Martin Danelljan*
- 18. Learning Continuous Implicit Representation for Near-Periodic Patterns, *Bowei Chen, Tiancheng Zhi, Martial Hebert, Srinivasa G. Narasimhan*
- 19. Improving the Reliability for Confidence Estimation, *Haoyuan Qu, Yanchao Li, Lin Geng Foo, Jason Kuen, Jiuxiang Gu, Jun Liu*
- 20. GraphVid: It Only Takes a Few Nodes to Understand a Video, *Eitan Kosman, Dotan Di Castro*
- 21. Towards Robust Face Recognition with Comprehensive Search, *Manyuan Zhang, Guanglu Song, Yu Liu, Hongsheng Li*
- 22. Pose2Room: Understanding 3D Scenes from Human Activities, *Yinyu Nie, Angela Dai, Xiaoguang Han, Matthias Nießner*
- 23. S₂Contact: Graph-Based Network for 3D Hand-Object Contact Estimation with Semi-Supervised Learning, *Tze Ho Elden Tse, Zhongqun Zhang, Kwang In Kim, Aleš Leonardis, Feng Zheng, Hyung Jin Chang*

24. DynaST: Dynamic Sparse Transformer for Exemplar-Guided Image Generation, *Songhua Liu, Jingwen Ye, Sucheng Ren, Xinchao Wang*
25. TAFIM: Targeted Adversarial Attacks against Facial Image Manipulations, *Shivangi Aneja, Lev Markhasin, Matthias Nießner*
26. EdgeViTs: Competing Light-Weight CNNs on Mobile Devices with Vision Transformers, *Junting Pan, Adrian Bulat, Fuwen Tan, Xiatian Zhu, Lukasz Dudziak, Hongsheng Li, Georgios Tzimiropoulos, Brais Martinez*
27. TokenMix: Rethinking Image Mixing for Data Augmentation in Vision Transformers, *Jihao Liu, Boxiao Liu, Hang Zhou, Hongsheng Li, Yu Liu*
28. Master of All: Simultaneous Generalization of Urban-Scene Segmentation to All Adverse Weather Conditions, *Nikhil Reddy, Abhinav Singhal, Abhishek Kumar, Mahsa Baktashmotlagh, Chetan Arora*
29. DFNet: Enhance Absolute Pose Regression with Direct Feature Matching, *Shuai Chen, Xinghui Li, Zirui Wang, Victor Adrian Prisacariu*
30. PrivHAR: Recognizing Human Actions from Privacy-Preserving Lens, *Carlos Hinojosa, Miguel Marquez, Henry Arguello, Ehsan Adeli, Li Fei-Fei, Juan Carlos Nibbles*
31. MeshUDF: Fast and Differentiable Meshing of Unsigned Distance Field Networks, *Benoît Guillard, Federico Stella, Pascal Fua*
32. Learn2Augment: Learning to Composite Videos for Data Augmentation in Action Recognition, *Shreyank N Gowda, Marcus Rohrbach, Frank Keller, Laura Sevilla-Lara*
33. Depth Map Decomposition for Monocular Depth Estimation, *Jinyoung Jun, Jae-Han Lee, Chul Lee, Chang-Su Kim*
34. Video Extrapolation in Space and Time, *Yunzhi Zhang, Jiajun Wu*
35. Learning Visual Styles from Audio-Visual Associations, *Tingle Li, Yichen Liu, Andrew Owens, Hang Zhao*
36. Learning Mutual Modulation for Self-Supervised Cross-Modal Super-Resolution, *Xiaoyu Dong, Naoto Yokoya, Longguang Wang, Tatsumi Uezato*
37. DeMFI: Deep Joint Deblurring and Multi-Frame Interpolation with Flow-Guided Attentive Correlation and Recursive Boosting, *Jihyong Oh, Munchurl Kim*
38. Sliced Recursive Transformer, *Zhiqiang Shen, Zechun Liu, Eric Xing*
39. UniNet: Unified Architecture Search with Convolution, Transformer, and MLP, *Jihao Liu, Xin Huang, Guanglu Song, Hongsheng Li, Yu Liu*
40. Flow Graph to Video Grounding for Weakly-Supervised Multi-step Localization, *Nikita Dvornik, Isma Hadji, Hai Pham, Dhaivat Bhatt, Brais Martinez, Afsaneh Fazly, Allan D. Jepson*
41. Facial Depth and Normal Estimation Using Single Dual-Pixel Camera, *Minjun Kang, Jaesung Choe, Hyowon Ha, Hae-Gon Jeon, Sunghoon Im, In So Kweon, Kuk-Jin Yoon*
42. Neural Scene Decoration from a Single Photograph, *Hong-Wing Pang, Yingshu Chen, Phuoc-Hieu Le, Binh-Son Hua, Thanh Nguyen, Sai-Kit Yeung*
43. Hallucinating Pose-Compatible Scenes, *Tim Brooks, Alexei A. Efros*
44. Few 'Zero Level Set'-Shot Learning of Shape Signed Distance Functions in Feature Space, *Amine Ouasfi, Adnane Boukhayma*
45. diffConv: Analyzing Irregular Point Clouds with an Irregular View, *Manxi Lin, Aasa Feragen*
46. TACS: Taxonomy Adaptive Cross-Domain Semantic Segmentation, *Rui Gong, Martin Danelljan, Dengxin Dai, Danda Pani Paudel, Ajad Chhatkuli, Fisher Yu, Luc Van Gool*
47. Weight Fixing Networks, *Christopher Subia-Waud, Srinandan Dasmahapatra*
48. Directed Ray Distance Functions for 3D Scene Reconstruction, *Nilesh Kulkarni, Justin Johnson, David F. Fouhey*
49. FS-COCO: Towards Understanding of Freehand Sketches of Common Objects in Context, *Pinaki Nath Chowdhury, Aneeshan Sain, Ayan Kumar Bhunia, Tao Xiang, Yulia Gryaditskaya, Yi-Zhe Song*
50. Bi-PointFlowNet: Bidirectional Learning for Point Cloud Based Scene Flow Estimation, *Wencan Cheng, Jong Hwan Ko*
51. A Dataset for Interactive Vision-Language Navigation with Unknown Command Feasibility, *Andrea Burns, Deniz Arsan, Sanjna Agrawal, Ranjitha Kumar, Kate Saenko, Bryan A. Plummer*
52. PSS: Progressive Sample Selection for Open-World Visual Representation Learning, *Tianyue Cao, Yongxin Wang, Yifan Xing, Tianjun Xiao, Tong He, Zheng Zhang, Hao Zhou, Joseph Tighe*
53. Point MixSwap: Attentional Point Cloud Mixing via Swapping Matched Structural Divisions, *Ardian Umam, Cheng-Kun Yang, Yung-Yu Chuang, Jen-Hui Chuang, Yen-Yu Lin*
54. RayTran: 3D Pose Estimation and Shape Reconstruction of Multiple Objects from Videos with Ray-Traced Transformers, *Michał J. Tyszkiewicz, Kevis-Kokitsi Maninis, Stefan Popov, Vittorio Ferrari*
55. Space-Partitioning RANSAC, *Daniel Barath, Gábor Valasek*
56. SPIN: An Empirical Evaluation on Sharing Parameters of Isotropic Networks, *Chien-Yu Lin, Anish Prabhu, Thomas Merth, Sachin Mehta, Anurag Ranjan, Maxwell Horton, Mohammad Rastegari*
57. Controllable Shadow Generation Using Pixel Height Maps, *Yichen Sheng, Yifan Liu, Jianming Zhang, Wei Yin, A. Cengiz Oztireli, He Zhang, Zhe Lin, Eli Shechtman, Bedrich Benes*
58. Contextformer: A Transformer with Spatio-Channel Attention for Context Modeling in Learned Image Compression, *A. Burakhan Koyuncu, Han Gao, Atanas Boev, Georgii Gaikov, Elena Alshina, Eckehard Steinbach*
59. MUGEN: A Playground for Video-Audio-Text Multimodal Understanding and GENeration, *Thomas Hayes, Songyang Zhang, Xi Yin, Guan Pang, Sasha Sheng, Harry Yang, Songwei Ge, Qiyuan Hu, Devi Parikh*
60. Bayesian Tracking of Video Graphs Using Joint Kalman Smoothing and Registration, *Aditi Basu Bal, Ramy Mounir, Sathyanarayanan Aakur, Sudeep Sarkar, Anuj Srivastava*
61. Video Dialog As Conversation about Objects Living in Space-Time, *Hoang-Anh Pham, Thao Minh Le, Vuong Le, Tu Minh Phuong, Truyen Tran*
62. Adversarial Label Poisoning Attack on Graph Neural Networks via Label Propagation, *Ganlin Liu, Xiaowei Huang, Xinping Yi*
63. Visual Knowledge Tracing, *Neehar Kondapaneni, Pietro Perona, Oisin Mac Aodha*
64. FedLTN: Federated Learning for Sparse and Personalized Lottery Ticket Networks, *Vaikkunth Mugunthan, Eric Lin, Vignesh Gokul, Christian Lau, Lalana Kagal, Steve Pieper*
65. Improving the Perceptual Quality of 2D Animation Interpolation, *Shuhong Chen, Matthias Zwicker*

66. MTTrans: Cross-Domain Object Detection with Mean Teacher Transformer, *Jinze Yu, Jiaming Liu, Xiaobao Wei, Haoyi Zhou, Yohei Nakata, Denis Gudovskiy, Tomoyuki Okuno, Jianxin Li, Kurt Keutzer, Shanghang Zhang*
67. Contrastive Learning for Diverse Disentangled Foreground Generation, *Yuheng Li, Yijun Li, Jingwan Lu, Eli Shechtman, Yong Jae Lee, Krishna Kumar Singh*
68. Fabric Material Recovery from Video Using Multi-Scale Geometric Auto-Encoder, *Junbang Liang, Ming Lin*
69. DSR – A Dual Subspace Re-Projection Network for Surface Anomaly Detection, *Vitjan Zavrtanik, Matej Kristan, Danijel Skočaj*
70. Perceptual Artifacts Localization for Inpainting, *Lingzhi Zhang, Yuqian Zhou, Connelly Barnes, Sohrab Amirghodsi, Zhe Lin, Eli Shechtman, Jianbo Shi*
71. ExtrudeNet: Unsupervised Inverse Sketch-and-Extrude for Shape Parsing, *Daxuan Ren, Jianmin Zheng, Jianfei Cai, Jiatong Li, Junzhe Zhang*
72. Demystifying Unsupervised Semantic Correspondence Estimation, *Mehmet Aygün, Oisín Mac Aodha*
73. LiDAL: Inter-Frame Uncertainty Based Active Learning for 3D LiDAR Semantic Segmentation, *Zeyu Hu, Xuyang Bai, Runze Zhang, Xin Wang, Guangyuan Sun, Hongbo Fu, Chiew-Lan Tai*
74. FashionViL: Fashion-Focused Vision-and-Language Representation Learning, *Xiao Han, Licheng Yu, Xiatian Zhu, Li Zhang, Yi-Zhe Song, Tao Xiang*
75. CycDA: Unsupervised Cycle Domain Adaptation to Learn from Image to Video, *Wei Lin, Anna Kukleva, Kunyang Sun, Horst Possegger, Hilde Kuehne, Horst Bischof*
76. NeLLF: Neural Incident Light Field for Physically-Based Material Estimation, *Yao Yao, Jingyang Zhang, Jingbo Liu, Yihang Qu, Tian Fang, David McKinnon, Yanghai Tsin, Long Quan*
77. What to Hide from Your Students: Attention-Guided Masked Image Modeling, *Ioannis Kakogeorgiou, Spyros Gidaris, Bill Psomas, Yannis Avrithis, Andrei Bursuc, Konstantinos Karantzas, Nikos Komodakis*
78. CPO: Change Robust Panorama to Point Cloud Localization, *Junho Kim, Hojun Jang, Changwoon Choi, Young Min Kim*
79. Streaming Multiscale Deep Equilibrium Models, *Can Ufuk Ertenli, Emre Akbas, Ramazan Gokberk Cinbis*
80. Secrets of Event-Based Optical Flow, *Shintaro Shiba, Yoshimitsu Aoki, Guillermo Gallego*
81. Unified Fully and Timestamp Supervised Temporal Action Segmentation via Sequence-to-Sequence Translation, *Nadine Behrmann, S. Alireza Golestaneh, Zico Kolter, Jürgen Gall, Mehdi Noroozi*
82. UnrealEgo: A New Dataset for Robust Egocentric 3D Human Motion Capture, *Hiroyasu Akada, Jian Wang, Soshi Shimada, Masaki Takahashi, Christian Theobalt, Vladislav Golyanik*
83. Augmentation of rPPG Benchmark Datasets: Learning to Remove and Embed rPPG Signals via Double Cycle Consistent Learning from Unpaired Facial Videos, *Cheng-Ju Hsieh, Wei-Hao Chung, Chiou-Ting Hsu*
84. Learning Instance and Task-Aware Dynamic Kernels for Few-Shot Learning, *Rongkai Ma, Pengfei Fang, Gil Avraham, Yan Zuo, Tianyu Zhu, Tom Drummond, Mehrtash Harandi*
85. Neural Image Representations for Multi-Image Fusion and Layer Separation, *Seonghyeon Nam, Marcus A. Brubaker, Michael S. Brown*
86. Neural Video Compression Using GANs for Detail Synthesis and Propagation, *Fabian Mentzer, Eirikur Agustsson, Johannes Ballé, David Minnen, Nick Johnston, George Toderici*
87. Camera Pose Estimation and Localization with Active Audio Sensing, *Karren Yang, Michael Firman, Eric Brachmann, Clément Godard*
88. HULC: 3D HUman Motion Capture with Pose Manifold SampLing and Dense Contact Guidance, *Soshi Shimada, Vladislav Golyanik, Zhi Li, Patrick Pérez, Weipeng Xu, Christian Theobalt*
89. Audio-Visual Mismatch-Aware Video Retrieval via Association and Adjustment, *Sangmin Lee, Sungjune Park, Yong Man Ro*
90. Uncertainty-DTW for Time Series and Sequences, *Lei Wang, Piotr Koniusz*
91. Contrastive Vicinal Space for Unsupervised Domain Adaptation, *Jaemin Na, Dongyoon Han, Hyung Jin Chang, Wonjun Hwang*
92. GRIT-VLP: Grouped Mini-Batch Sampling for Efficient Vision and Language Pre-training, *Jaeseok Byun, Taebaek Hwang, Jianlong Fu, Taesup Moon*
93. Helpful or Harmful: Inter-Task Association in Continual Learning, *Hyundong Jin, Eunwoo Kim*
94. SAGA: Stochastic Whole-Body Grasping with Contact, *Yan Wu, Jiahao Wang, Yan Zhang, Siwei Zhang, Otmar Hilliges, Fisher Yu, Siyu Tang*
95. Style Your Hair: Latent Optimization for Pose-Invariant Hairstyle Transfer via Local-Style-Aware Hair Alignment, *Taewoo Kim, Chaeyeon Chung, Yoonseo Kim, Sunghyun Park, Kangyeol Kim, Jaegul Choo*
96. NashAE: Disentangling Representations through Adversarial Covariance Minimization, *Eric Yeats, Frank Liu, David Womble, Hai Li*
97. ManiFest: Manifold Deformation for Few-Shot Image Translation, *Fabio Pizzati, Jean-François Lalonde, Raoul de Charette*
98. Improving Generalization in Federated Learning by Seeking Flat Minima, *Debora Caldarola, Barbara Caputo, Marco Ciccone*
99. MultiMAE: Multi-modal Multi-task Masked Autoencoders, *Roman Bachmann, David Mizrahi, Andrei Atanov, Amir Zamir*
100. InfiniteNature-Zero: Learning Perpetual View Generation of Natural Scenes from Single Images, *Zhengqi Li, Qianqian Wang, Noah Snavely, Angjoo Kanazawa*
101. PoserNet: Refining Relative Camera Poses Exploiting Object Detections, *Matteo Taiana, Matteo Toso, Stuart James, Alessio Del Bue*
102. Autoregressive Uncertainty Modeling for 3D Bounding Box Prediction, *YuXuan Liu, Nikhil Mishra, Maximilian Sieb, Yide Shentu, Pieter Abbeel, Xi Chen*
103. VizWiz-FewShot: Locating Objects in Images Taken by People with Visual Impairments, *Yu-Yun Tseng, Alexander Bell, Danna Gurari*
104. Transfer without Forgetting, *Matteo Boschini, Lorenzo Bonicelli, Angelo Porrello, Giovanni Bellitto, Matteo Pennisi, Simone Palazzo, Concetto Spampinato, Simone Calderara*
105. Multiview Stereo with Cascaded Epipolar RAFT, *Zeyu Ma, Zachary Teed, Jia Deng*
106. Generative Adversarial Network for Future Hand Segmentation from Egocentric Video, *Wenqi Jia, Miao Liu, James M. Rehg*

-
- A full page of blank graph paper. The grid consists of 20 columns and 20 rows of small squares, formed by thin black lines. There are no margins or other markings on the page.

[illegible][illegible][illegible][illegible][illegible][illegible]A full-page sheet of white graph paper with a light gray grid. The grid consists of small squares, approximately 10 units wide by 10 units high. There are no margins or additional markings on the page.

Thursday, October 27

0700–1830 Registration (Pavillion 2 Lobby)

0800–0900 Coffee & Refreshments (Lobby and Hall C)

0800–0900 Poster Setup (Hall B)

0900–1630 Exhibits (Hall C)

- See online for exhibitor list and map.

0900–1530 Demos (Hall C Demo Area)

- Video from Coded Motion Blur Using Dynamic Phase Coding, *Erez Yosef, Shay Elmalem, Raja Giryes (Tel Aviv Univ., Israel)*
- Using a Smartphone for Augmented Reality in a Classroom (Two Applications), *Yair Moshe, Dan-Ilán Ben-David, Eran Mann, Sagie Baboach, Iddo Bar-Haim, Shoval Gerbi, Adam Katav (Technion - Israel Institute of Technology, Israel)*
- EMScan: A Mobile Application for Early Lyme Disease Diagnosis, *Yann FRENDO, Pr. E. MEPHU-NGUIFO (Université Clermont Auvergne, France)*
- Learning Phase Mask for Privacy-Preserving Passive Depth Estimation, *Francesco Pittaluga (NEC Labs America, USA), Zaid Tasneem (Rice Univ.)*

0900–1100 Oral 3.A.1: Stereo and 3D Multiview/ Sensors (Hall D; overflow in Halls K-M)

Papers in this session are in Poster Session 3.A

Format (9 min. presentation; 3 min. questions)

- [0900] Box2Mask: Weakly Supervised 3D Semantic Instance Segmentation Using Bounding Boxes, *Julian Chibane, Francis Engelmann, Tuan Anh Tran, Gerard Pons-Moll*
- [0912] Generalizable Patch-Based Neural Rendering, *Mohammed Suhail, Carlos Esteves, Leonid Sigal, Ameesh Makadia*
- [0924] Solution Space Analysis of Essential Matrix Based on Algebraic Error Minimization, *Gaku Nakano*
- [0936] Approximate Differentiable Rendering with Algebraic Surfaces, *Leonid Keselman, Martial Hebert*
- [0948] Beyond Periodicity: Towards a Unifying Framework for Activations in Coordinate-MLPs, *Sameera Ramasinghe, Simon Lucey*
- [1000] Gaussian Activated Neural Radiance Fields for High Fidelity Reconstruction & Pose Estimation, *Shin-Fang Chng, Sameera Ramasinghe, Jamie Sherrah, Simon Lucey*
- [1012] Unbiased Gradient Estimation for Differentiable Surface Splatting via Poisson Sampling, *Jan U. Müller, Michael Weinmann, Reinhard Klein*

0900–1100 Oral 3.A.2: Datasets & Evaluation (Hall E; overflow in Halls F-H)

Papers in this session are in Poster Session 3.A

Format (9 min. presentation; 3 min. questions)

- [0900] OccamNets: Mitigating Dataset Bias by Favoring Simpler Hypotheses, *Robik Shrestha, Kushal Kafle, Christopher Kanan*
- [0912] The Abduction of Sherlock Holmes: A Dataset for Visual Abductive Reasoning, *Jack Hessel, Jena D. Hwang, Jae Sung Park, Rowan Zellers, Chandra Bhagavatula, Anna Rohrbach, Kate Saenko, Yejin Choi*

- [0924] Look Both Ways: Self-Supervising Driver Gaze Estimation and Road Scene Saliency, *Isaac Kasahara, Simon Stent, Hyun Soo Park*
- [0936] A Dense Material Segmentation Dataset for Indoor and Outdoor Scene Parsing, *Paul Upchurch, Ransen Niu*
- [0948] "This Is My Unicorn, Fluffy": Personalizing Frozen Vision-Language Representations, *Niv Cohen, Rinon Gal, Eli A. Meirum, Gal Chechik, Yuval Atzmon*
- [1000] HuMMan: Multi-modal 4D Human Dataset for Versatile Sensing and Modeling, *Zhongang Cai, Daxuan Ren, Ailing Zeng, Zhengyu Lin, Tao Yu, Wenjia Wang, Xiangyu Fan, Yang Gao, Yifan Yu, Liang Pan, Fangzhou Hong, Mingyuan Zhang, Chen Change Loy, Lei Yang, Ziwei Liu*

1100–1200 Coffee & Refreshments (Lobby and Hall C)

1100–1330 Poster 3.A (Hall B)

1. WeLSA: Learning to Predict 6D Pose from Weakly Labeled Data Using Shape Alignment, *Shishir Reddy Vutukur, Ivan Shugurov, Benjamin Busam, Andreas Hutter, Slobodan Ilic*
2. BEAT: A Large-Scale Semantic and Emotional Multi-modal Dataset for Conversational Gestures Synthesis, *Haiyang Liu, Zihao Zhu, Naoya Iwamoto, Yichen Peng, Zhengqing Li, You Zhou, Elif Bozkurt, Bo Zheng*
3. Implicit Neural Representations for Image Compression, *Yannick Strümpfer, Janis Postels, Ren Yang, Luc Van Gool, Federico Tombari*
4. Neural Architecture Search for Spiking Neural Networks, *Youngeun Kim, Yuhang Li, Hyoungseob Park, Yeshwanth Venkatesha, Priyadarshini Panda*
5. Semi-Supervised Monocular 3D Object Detection by Multi-View Consistency, *Qing Lian, Yanbo Xu, Weilong Yao, Yingcong Chen, Tong Zhang*
6. TALISMAN: Targeted Active Learning for Object Detection with Rare Classes and Slices Using Submodular Mutual Information, *Suraj Kothawade, Saikat Ghosh, Sumit Shekhar, Yu Xiang, Rishabh Iyer*
7. S2F2: Single-Stage Flow Forecasting for Future Multiple Trajectories Prediction, *Yu-Wen Chen, Hsuan-Kung Yang, Chu-Chi Chiu, Chun-Yi Lee*
8. 3D Compositional Zero-Shot Learning with DeCompositional Consensus, *Muhammad Ferjad Naeem, Evin Pinar Örnek, Yongqin Xian, Luc Van Gool, Federico Tombari*
9. Real-Time Online Video Detection with Temporal Smoothing Transformers, *Yue Zhao, Philipp Krähenbühl*
10. Box2Mask: Weakly Supervised 3D Semantic Instance Segmentation Using Bounding Boxes, *Julian Chibane, Francis Engelmann, Tuan Anh Tran, Gerard Pons-Moll*
11. Learning to Drive by Watching YouTube Videos: Action-Conditioned Contrastive Policy Pretraining, *Qihang Zhang, Zhenghao Peng, Bolei Zhou*
12. Making Heads or Tails: Towards Semantically Consistent Visual Counterfactuals, *Simon Vandenhende, Dhruv Mahajan, Filip Radenovic, Deepti Ghadiyaram*
13. Towards Generic 3D Tracking in RGBD Videos: Benchmark and Baseline, *Jinyu Yang, Zhongqun Zhang, Zhe Li, Hyung Jin Chang, Aleš Leonardis, Feng Zheng*
14. ViewFormer: NeRF-Free Neural Rendering from Few Images Using Transformers, *Jonáš Kulháněk, Erik Derner, Torsten Sattler, Robert Babuška*

15. Generalizable Patch-Based Neural Rendering, *Mohammed Suhail, Carlos Esteves, Leonid Sigal, Ameesh Makadia*
16. Not Just Streaks: Towards Ground Truth for Single Image Deraining, *Yunhao Ba, Howard Zhang, Ethan Yang, Akira Suzuki, Arnold Pfahnl, Chethan Chinder Chandrappa, Celso M. de Melo, Suyu You, Stefano Soatto, Alex Wong, Achuta Kadambi*
17. Self-supervised Human Mesh Recovery with Cross-Representation Alignment, *Xuan Gong, Meng Zheng, Benjamin Planche, Srikrishna Karanam, Terrence Chen, David Doermann, Ziyang Wu*
18. Neural Density-Distance Fields, *Itsuki Ueda, Yoshihiro Fukuhara, Hirokatsu Kataoka, Hiroaki Aizawa, Hidehiko Shishido, Itaru Kitahara*
19. BungeeNeRF: Progressive Neural Radiance Field for Extreme Multi-Scale Scene Rendering, *Yuanbo Xiangli, Linning Xu, Xingang Pan, Nanxuan Zhao, Anyi Rao, Christian Theobalt, Bo Dai, Dahua Lin*
20. Solution Space Analysis of Essential Matrix Based on Algebraic Error Minimization, *Gaku Nakano*
21. SESS: Saliency Enhancing with Scaling and Sliding, *Osman Tursun, Simon Denman, Sridha Sridharan, Clinton Fookes*
22. Detecting Twenty-Thousand Classes Using Image-Level Supervision, *Xingyi Zhou, Rohit Girdhar, Armand Joulin, Philipp Krähenbühl, Ishan Misra*
23. Max Pooling with Vision Transformers Reconciles Class and Shape in Weakly Supervised Semantic Segmentation, *Simone Rossetti, Damiano Zappia, Marta Sanzari, Marco Schaerf, Fiora Pirri*
24. Context-Aware Streaming Perception in Dynamic Environments, *Gur-Eyal Sela, Ionel Gog, Justin Wong, Kumar Krishna Agrawal, Xiangxi Mo, Sukrit Kalra, Peter Schafhalter, Eric Leong, Xin Wang, Bharathan Balaji, Joseph Gonzalez, Ion Stoica*
25. Approximate Differentiable Rendering with Algebraic Surfaces, *Leonid Keselman, Martial Hebert*
26. Doubly Deformable Aggregation of Covariance Matrices for Few-Shot Segmentation, *Zhitong Xiong, Haopeng Li, Xiao Xiang Zhu*
27. MemSAC: Memory Augmented Sample Consistency for Large Scale Domain Adaptation, *Tarun Kalluri, Astuti Sharma, Manmohan Chandraker*
28. Efficient One-Stage Video Object Detection by Exploiting Temporal Consistency, *Guanxiong Sun, Yang Hua, Guosheng Hu, Neil Robertson*
29. Domain Generalization by Mutual-Information Regularization with Pre-trained Models, *Junbum Cha, Kyungjae Lee, Sungrae Park, Sanghyuk Chun*
30. Beyond Periodicity: Towards a Unifying Framework for Activations in Coordinate-MLPs, *Sameera Ramasinghe, Simon Lucey*
31. SNeS: Learning Probably Symmetric Neural Surfaces from Incomplete Data, *Eldar Insafutdinov, Dylan Campbell, João F. Henriques, Andrea Vedaldi*
32. HDR-Plenoxels: Self-Calibrating High Dynamic Range Radiance Fields, *Kim Jun-Seong, Kim Yu-Ji, Moon Ye-Bin, Tae-Hyun Oh*
33. Abstracting Sketches through Simple Primitives, *Stephan Alaniz, Massimiliano Mancini, Anjan Dutta, Diego Marcos, Zeynep Akata*
34. Adversarial Contrastive Learning via Asymmetric InfoNCE, *Qiyang Yu, Jieming Lou, Xianyu Zhan, Qizhang Li, Wangmeng Zuo, Yang Liu, Jingjing Liu*
35. A-OKVQA: A Benchmark for Visual Question Answering Using World Knowledge, *Dustin Schwenk, Apoorv Khandelwal, Christopher Clark, Kenneth Marino, Roozbeh Mottaghi*
36. Relationformer: A Unified Framework for Image-to-Graph Generation, *Suprosanna Shit, Rajat Koner, Bastian Wittmann, Johannes Paetzold, Ivan Ezhov, Hongwei Li, Jiazhen Pan, Sahand Sharifzadeh, Georgios Kaissis, Volker Tresp, Bjoern Menze*
37. SLIP: Self-Supervision Meets Language-Image Pre-training, *Norman Mu, Alexander Kirillov, David Wagner, Saining Xie*
38. IS-MVSNet: Importance Sampling-Based MVSNet, *Likang Wang, Yue Gong, Xinjun Ma, Qirui Wang, Kaixuan Zhou, Lei Chen*
39. PointMixer: MLP-Mixer for Point Cloud Understanding, *Jaesung Choe, Chunghyun Park, Francois Rameau, Jaesik Park, In So Kweon*
40. Gaussian Activated Neural Radiance Fields for High Fidelity Reconstruction & Pose Estimation, *Shin-Fang Chng, Sameera Ramasinghe, Jamie Sherrah, Simon Lucey*
41. Perspective Flow Aggregation for Data-Limited 6D Object Pose Estimation, *Yinlin Hu, Pascal Fua, Mathieu Salzmann*
42. ChunkyGAN: Real Image Inversion via Segments, *Adéla Šubrtová, David Futschik, Jan Čech, Michal Lukáč, Eli Shechtman, Daniel Šýkora*
43. 3D-PL: Domain Adaptive Depth Estimation with 3D-Aware Pseudo-Labeling, *Yu-Ting Yen, Chia-Ni Lu, Wei-Chen Chiu, Yi-Hsuan Tsai*
44. Camera Pose Auto-Encoders for Improving Pose Regression, *Yoli Shavit, Yosi Keller*
45. FindIt: Generalized Localization with Natural Language Queries, *Weicheng Kuo, Fred Bertsch, Wei Li, AJ Piergiovanni, Mohammad Saffar, Anelia Angelova*
46. Action-Based Contrastive Learning for Trajectory Prediction, *Marah Halawa, Olaf Hellwich, Pia Bideau*
47. TDViT: Temporal Dilated Video Transformer for Dense Video Tasks, *Guanxiong Sun, Yang Hua, Guosheng Hu, Neil Robertson*
48. SimpleRecon: 3D Reconstruction without 3D Convolutions, *Mohamed Sayed, John Gibson, Jamie Watson, Victor Prisacariu, Michael Firman, Clément Godard*
49. CADyQ: Content-Aware Dynamic Quantization for Image Super-Resolution, *Cheeun Hong, Sungyong Baik, Heewon Kim, Seungjun Nah, Kyoung Mu Lee*
50. Unbiased Gradient Estimation for Differentiable Surface Splatting via Poisson Sampling, *Jan U. Müller, Michael Weinmann, Reinhard Klein*
51. A Kendall Shape Space Approach to 3D Shape Estimation from 2D Landmarks, *Martha Paskin, Daniel Baum, Mason N. Dean, Christoph von Tycowicz*
52. Pre-training Strategies and Datasets for Facial Representation Learning, *Adrian Bulat, Shiyang Cheng, Jing Yang, Andrew Garbett, Enrique Sanchez, Georgios Tzimiropoulos*
53. Neural Strands: Learning Hair Geometry and Appearance from Multi-View Images, *Radu Alexandru Rosu, Shunsuke Saito, Ziyang Wang, Chenglei Wu, Sven Behnke, Giljoo Nam*

54. Improving Few-Shot Learning through Multi-task Representation Learning Theory, *Quentin Bouniot, Ievgen Redko, Romaric Audigier, Angélique Loesch, Amaury Habrard*
55. Long-Tailed Instance Segmentation Using Gumbel Optimized Loss, *Konstantinos Panagiotis Alexandridis, Jiankang Deng, Anh Nguyen, Shan Luo*
56. 3D Scene Inference from Transient Histograms, *Sacha Jungerman, Atul Ingle, Yin Li, Mohit Gupta*
57. Network Binarization via Contrastive Learning, *Yuzhang Shang, Dan Xu, Ziliang Zong, Liqiang Nie, Yan Yan*
58. Is Geometry Enough for Matching in Visual Localization? *Qunjie Zhou, Sérgio Agostinho, Aljoša Ošep, Laura Leal-Taixé*
59. 3D Face Reconstruction with Dense Landmarks, *Erroll Wood, Tadas Baltrušaitis, Charlie Hewitt, Matthew Johnson, Jingjing Shen, Nikola Milosavljević, Daniel Wilde, Stephan Garbin, Toby Sharp, Ivan Stojiljković, Tom Cashman, Julien Valentin*
60. OccamNets: Mitigating Dataset Bias by Favoring Simpler Hypotheses, *Robik Shrestha, Kushal Kafle, Christopher Kanan*
61. GLAMD: Global and Local Attention Mask Distillation for Object Detectors, *Younho Jang, Wheemyung Shin, Jinbeom Kim, Simon Woo, Sung-Ho Bae*
62. Language-Grounded Indoor 3D Semantic Segmentation in the Wild, *Dávid Rozenberszki, Or Litany, Angela Dai*
63. Entry-Flipped Transformer for Inference and Prediction of Participant Behavior, *Bo Hu, Tat-Jen Cham*
64. DoodleFormer: Creative Sketch Drawing with Transformers, *Ankan Kumar Bhunia, Salman Khan, Hisham Cholakkal, Rao Muhammad Anwer, Fahad Shahbaz Khan, Jorma Laaksonen, Michael Felsberg*
65. Compositional Visual Generation with Composable Diffusion Models, *Nan Liu, Shuang Li, Yilun Du, Antonio Torralba, Joshua B. Tenenbaum*
66. 3D Shape Sequence of Human Comparison and Classification Using Current and Varifolds, *Emery Pierson, Mohamed Daoudi, Sylvain Arguillere*
67. FEAR: Fast, Efficient, Accurate and Robust Visual Tracker, *Vasyl Borsuk, Roman Vei, Orest Kupyn, Tetiana Martyniuk, Igor Krasheniy, Jiří Matas*
68. Learning Phase Mask for Privacy-Preserving Passive Depth Estimation, *Zaid Tasneem, Giovanni Milione, Yi-Hsuan Tsai, Xiang Yu, Ashok Veeraraghavan, Manmohan Chandraker, Francesco Pittaluga*
69. Break and Make: Interactive Structural Understanding Using LEGO Bricks, *Aaron Walsman, Muru Zhang, Klemen Kotar, Karthik Desingh, Ali Farhadi, Dieter Fox*
70. Event-Based Fusion for Motion Deblurring with Cross-Modal Attention, *Lei Sun, Christos Sakaridis, Jingyun Liang, Qi Jiang, Kailun Yang, Peng Sun, Yaozu Ye, Kaiwei Wang, Luc Van Gool*
71. Pure Transformer with Integrated Experts for Scene Text Recognition, *Yew Lee Tan, Adams Wai-Kin Kong, Jung-Jae Kim*
72. A Simple Single-Scale Vision Transformer for Object Detection and Instance Segmentation, *Wuyang Chen, Xianzhi Du, Fan Yang, Lucas Beyer, Xiaohua Zhai, Tsung-Yi Lin, Huizhong Chen, Jing Li, Xiaodan Song, Zhangyang Wang, Denny Zhou*
73. Generating Natural Images with Direct Patch Distributions Matching, *Ariel Elnekave, Yair Weiss*
74. TRoVE: Transforming Road Scene Datasets into Photorealistic Virtual Environments, *Shubham Dokania, Anbumani Subramanian, Manmohan Chandraker, C.V. Jawahar*
75. RDO-Q: Extremely Fine-Grained Channel-Wise Quantization via Rate-Distortion Optimization, *Zhe Wang, Jie Lin, Xue Geng, Mohamed M. Sabry Aly, Vijay Chandrasekhar*
76. Understanding Collapse in Non-Contrastive Siamese Representation Learning, *Alexander C. Li, Alexei A. Efros, Deepak Pathak*
77. Trading Positional Complexity vs Deepness in Coordinate Networks, *Jianqiao Zheng, Sameera Ramasinghe, Xueqian Li, Simon Lucey*
78. U-Boost NAS: Utilization-Boosted Differentiable Neural Architecture Search, *Ahmet Caner Yüzügüler, Nikolaos Dimitriadis, Pascal Frossard*
79. Trapped in Texture Bias? A Large Scale Comparison of Deep Instance Segmentation, *Johannes Theodoridis, Jessica Hofmann, Johannes Maucher, Andreas Schilling*
80. The Abduction of Sherlock Holmes: A Dataset for Visual Abductive Reasoning, *Jack Hessel, Jena D. Hwang, Jae Sung Park, Rowan Zellers, Chandra Bhagavatula, Anna Rohrbach, Kate Saenko, Yejin Choi*
81. Object-Compositional Neural Implicit Surfaces, *Qianyi Wu, Xian Liu, Yuedong Chen, Kejie Li, Chuanxia Zheng, Jianfei Cai, Jianmin Zheng*
82. Sem2NeRF: Converting Single-View Semantic Masks to Neural Radiance Fields, *Yuedong Chen, Qianyi Wu, Chuanxia Zheng, Tat-Jen Cham, Jianfei Cai*
83. Unsupervised Night Image Enhancement: When Layer Decomposition Meets Light-Effects Suppression, *Yeying Jin, Wenhan Yang, Robby T. Tan*
84. Zero-Shot Temporal Action Detection via Vision-Language Prompting, *Sauradip Nag, Xiatian Zhu, Yi-Zhe Song, Tao Xiang*
85. Automatic Dense Annotation of Large-Vocabulary Sign Language Videos, *Liliane Momeni, Hannah Bull, K R Prajwal, Samuel Albanie, Gül Varol, Andrew Zisserman*
86. Making the Most of Text Semantics to Improve Biomedical Vision-Language Processing, *Benedikt Boecking, Naoto Usuyama, Shruthi Bannur, Daniel C. Castro, Anton Schwaighofer, Stephanie Hyland, Maria Wetscherek, Tristan Naumann, Aditya Nori, Javier Alvarez-Valle, Hoifung Poon, Ozan Oktay*
87. MOTCOM: The Multi-Object Tracking Dataset Complexity Metric, *Malte Pedersen, Joakim Bruslund Haurum, Patrick Dendorfer, Thomas B. Moeslund*
88. Feature Representation Learning for Unsupervised Cross-Domain Image Retrieval, *Conghui Hu, Gim Hee Lee*
89. A Broad Study of Pre-training for Domain Generalization and Adaptation, *Donghyun Kim, Kaihong Wang, Stan Sclaroff, Kate Saenko*
90. Look Both Ways: Self-Supervising Driver Gaze Estimation and Road Scene Saliency, *Isaac Kasahara, Simon Stent, Hyun Soo Park*
91. Improved Masked Image Generation with Token-Critic, *José Lezama, Huiwen Chang, Lu Jiang, Irfan Essa*
92. Semantic-Guided Multi-Mask Image Harmonization, *Xuqian Ren, Yifan Liu*
93. Object-Centric Unsupervised Image Captioning, *Zihang Meng, David Yang, Xuefei Cao, Ashish Shah, Ser-Nam Lim*
94. Event Neural Networks, *Matthew Dutson, Yin Li, Mohit Gupta*
95. A Non-Isotropic Probabilistic Take On Proxy-Based Deep Metric Learning, *Michael Kirchhof, Karsten Roth, Zeynep Akata, Enkelejda Kasneci*

- 1200-1330 Lunch** (Lobby and Hall C)

Notes:

[illegible]

1330–1430 Poster Switch/Setup (Hall B)

1330–1530 Oral 3.B.1: Robots, Vehicles, and Computational Photography (Hall D; overflow in Halls K-M)

Papers in this session are in Poster Session 3.B

Format (9 min. presentation; 3 min. questions)

- [1330] Practical and Scalable Desktop-Based High-Quality Facial Capture, *Alexandros Lattas, Yiming Lin, Jayanth Kannan, Ekin Ozturk, Luca Filipi, Giuseppe Claudio Guarnera, Gaurav Chawla, Abhijeet Ghosh*
- [1342] Drive&Segment: Unsupervised Semantic Segmentation of Urban Scenes via Cross-Modal Distillation, *Antonin Vobecky, David Hurych, Oriane Siméoni, Spyros Gidaris, Andrei Bursuc, Patrick Pérez, Josef Sivic*
- [1354] SpOT: Spatiotemporal Modeling for 3D Object Tracking, *Colton Stearns, Davis Rempe, Jie Li, Rareş Ambruş, Sergey Zakharov, Vitor Guizilini, Yanchao Yang, Leonidas J. Guibas*
- [1406] Synthesizing Light Field Video from Monocular Video, *Shrisudhan Govindarajan, Prasan Shedligeri, Sarah, Kaushik Mitra*
- [1418] LESS: Label-Efficient Semantic Segmentation for LiDAR Point Clouds, *Minghua Liu, Yin Zhou, Charles R. Qi, Boqing Gong, Hao Su, Dragomir Anguelov*
- [1430] EvAC3D: From Event-Based Apparent Contours to 3D Models via Continuous Visual Hulls, *Ziyun Wang, Kenneth Chaney, Kostas Daniilidis*
- [1442] Sim-2-Sim Transfer for Vision-and-Language Navigation in Continuous Environments, *Jacob Krantz, Stefan Lee*
- [1454] Pixel-Wise Energy-Biased Abstention Learning for Anomaly Segmentation on Complex Urban Driving Scenes, *Yu Tian, Yuyuan Liu, Guansong Pang, Fengbei Liu, Yuanhong Chen, Gustavo Carneiro*
- [1506] KING: Generating Safety-Critical Driving Scenarios for Robust Imitation via Kinematics Gradients, *Niklas Hanselmann, Katrin Renz, Kashyap Chitta, Apratim Bhattacharyya, Andreas Geiger*

1330–1530 Oral 3.B.2: Vision, Text, and Non-Supervised Learning (Hall E; overflow in Halls F-H)

Papers in this session are in Poster Session 3.B

Format (9 min. presentation; 3 min. questions)

- [1330] GLASS: Global to Local Attention for Scene-Text Spotting, *Roi Ronen, Shahar Tsiper, Oron Anshel, Inbal Lavi, Amir Markovitz, R. Manmatha*
- [1342] Registration Based Few-Shot Anomaly Detection, *Chaoqin Huang, Haoyan Guan, Aofan Jiang, Ya Zhang, Michael Spratling, Yan-Feng Wang*
- [1354] Decoupled Adversarial Contrastive Learning for Self-Supervised Adversarial Robustness, *Chaoning Zhang, Kang Zhang, Chenshuang Zhang, Axi Niu, Jiu Feng, Chang D. Yoo, In So Kweon*
- [1406] Weakly Supervised Grounding for VQA in Vision-Language Transformers, *Aisha Urooj, Hilde Kuehne, Chuang Gan, Niels Da Vitoria Lobo, Mubarak Shah*

1530–1630 Coffee & Refreshments (Lobby and Hall C)

1530–1730 Poster 3.B (Hall B)


1. Rethinking Generic Camera Models for Deep Single Image Camera Calibration to Recover Rotation and Fisheye Distortion, *Nobuhiko Wakai, Satoshi Sato, Yasunori Ishii, Takayoshi Yamashita*
2. COMPOSER: Compositional Reasoning of Group Activity in Videos with Keypoint-Only Modality, *Honglu Zhou, Asim Kadav, Aviv Shamsian, Shijie Geng, Farley Lai, Long Zhao, Ting Liu, Mubbasir Kapadia, Hans Peter Graf*
3. MvDeCor: Multi-View Dense Correspondence Learning for Fine-Grained 3D Segmentation, *Gopal Sharma, Kangxue Yin, Subhansu Maji, Evangelos Kalogerakis, Or Litany, Sanja Fidler*
4. Revisiting Point Cloud Simplification: A Learnable Feature Preserving Approach, *Rolandos Alexandros Potamias, Giorgos Bouritsas, Stefanos Zafeiriou*
5. GLASS: Global to Local Attention for Scene-Text Spotting, *Roi Ronen, Shahar Tsiper, Oron Anshel, Inbal Lavi, Amir Markovitz, R. Manmatha*
6. SOS! Self-Supervised Learning over Sets of Handled Objects in Egocentric Action Recognition, *Victor Escorcia, Ricardo Guerrero, Xiatian Zhu, Brais Martinez*
7. LaMAR: Benchmarking Localization and Mapping for Augmented Reality, *Paul-Edouard Sarlin, Mihai Dusmanu, Johannes L. Schönberger, Pablo Speciale, Lukas Gruber, Viktor Larsson, Ondrej Miksik, Marc Pollefeys*
8. Identity-Aware Hand Mesh Estimation and Personalization from RGB Images, *Deying Kong, Linguang Zhang, Liangjian Chen, Haoyu Ma, Xiangyi Yan, Shanlin Sun, Xingwei Liu, Kun Han, Xiaohui Xie*
9. Prune Your Model before Distill It, *Jinhyuk Park, Albert No*
10. Registration Based Few-Shot Anomaly Detection, *Chaoqin Huang, Haoyan Guan, Aofan Jiang, Ya Zhang, Michael Spratling, Yan-Feng Wang*
11. ASpanFormer: Detector-Free Image Matching with Adaptive Span Transformer, *Hongkai Chen, Zixin Luo, Lei Zhou, Yurun Tian, Mingmin Zhen, Tian Fang, David McKinnon, Yanghai Tsing, Long Quan*
12. Egocentric Activity Recognition and Localization on a 3D Map, *Miao Liu, Lingni Ma, Kiran Somasundaram, Yin Li, Kristen Grauman, James M. Rehg, Chao Li*
13. Video Question Answering with Iterative Video-Text Co-Tokenization, *AJ Piergiovanni, Kairo Morton, Weicheng Kuo, Michael S. Ryoo, Anelia Angelova*
14. LaTeRF: Label and Text Driven Object Radiance Fields, *Ashkan Mirzaei, Yash Kant, Jonathan Kelly, Igor Gilitschenski*
15. Decoupled Adversarial Contrastive Learning for Self-Supervised Adversarial Robustness, *Chaoning Zhang, Kang Zhang, Chenshuang Zhang, Axi Niu, Jiu Feng, Chang D. Yoo, In So Kweon*
16. CompNVS: Novel View Synthesis with Scene Completion, *Zuoyue Li, Tianxing Fan, Zhenqiang Li, Zhaopeng Cui, Yoichi Sato, Marc Pollefeys, Martin R. Oswald*
17. Fast Two-Step Blind Optical Aberration Correction, *Thomas Eboli, Jean-Michel Morel, Gabriele Facciolo*
18. PseudoClick: Interactive Image Segmentation with Click Imitation, *Qin Liu, Meng Zheng, Benjamin Planche, Srikrishna Karanam, Terrence Chen, Marc Niethammer, Ziyang Wu*
19. Trust, but Verify: Using Self-Supervised Probing to Improve Trustworthiness, *Ailin Deng, Shen Li, Miao Xiong, Zhirui Chen, Bryan Hooi*

20. Weakly Supervised Grounding for VQA in Vision-Language Transformers, *Aisha Urooj, Hilde Kuehne, Chuang Gan, Niels Da Vitoria Lobo, Mubarak Shah*
21. ASSISTER: Assistive Navigation via Conditional Instruction Generation, *Zanming Huang, Zhongkai Shanguan, Jimuyang Zhang, Gilad Bar, Matthew Boyd, Eshed Ohn-Bar*
22. Class-Agnostic Object Detection with Multi-modal Transformer, *Muhammad Maaz, Hanoona Rasheed, Salman Khan, Fahad Shahbaz Khan, Rao Muhammad Anwer, Ming-Hsuan Yang*
23. Learning an Isometric Surface Parameterization for Texture Unwrapping, *Sagnik Das, Ke Ma, Zhixin Shu, Dimitris Samaras*
24. Learning to Censor by Noisy Sampling, *Ayush Chopra, Abhinav Java, Abhishek Singh, Vivek Sharma, Ramesh Raskar*
25. TAVA: Template-Free Animatable Volumetric Actors, *Ruilong Li, Julian Tanke, Minh Vo, Michael Zollhöfer, Jürgen Gall, Angjoo Kanazawa, Christoph Lassner*
26. DEVIANT: Depth EquiVArLAnt NeTwork for Monocular 3D Object Detection, *Abhinav Kumar, Garrick Brazil, Enrique Corona, Armin Parchami, Xiaoming Liu*
27. Implicit Neural Representations for Variable Length Human Motion Generation, *Pablo Cervantes, Yusuke Sekikawa, Ikuro Sato, Koichi Shinoda*
28. Spatially Invariant Unsupervised 3D Object-Centric Learning and Scene Decomposition, *Tianyu Wang, Miaomiao Liu, Kee Siong Ng*
29. Negative Samples Are at Large: Leveraging Hard-Distance Elastic Loss for Re-identification, *Hyungtae Lee, Sungmin Eum, Heesung Kwon*
30. Practical and Scalable Desktop-Based High-Quality Facial Capture, *Alexandros Lattas, Yiming Lin, Jayanth Kannan, Ekin Ozturk, Luca Filipi, Giuseppe Claudio Guarnera, Gaurav Chawla, Abhijeet Ghosh*
31. FAR: Fourier Aerial Video Recognition, *Divya Kothandaraman, Tianrui Guan, Xijun Wang, Shuowen Hu, Ming Lin, Dinesh Manocha*
32. D3Net: A Unified Speaker-Listener Architecture for 3D Dense Captioning and Visual Grounding, *Zhenyu Chen, Qirui Wu, Matthias Nießner, Angel X. Chang*
33. D2ADA: Dynamic Density-Aware Active Domain Adaptation for Semantic Segmentation, *Tsung-Han Wu, Yi-Syuan Liou, Shao-Ji Yuan, Hsin-Ying Lee, Tung-I Chen, Kuan-Chih Huang, Winston H. Hsu*
34. FILM: Frame Interpolation for Large Motion, *Fitsum Reda, Janne Kontkanen, Eric Tabellion, Deqing Sun, Caroline Pantofaru, Brian Curless*
35. A Deep Moving-Camera Background Model, *Guy Erez, Ron Shapira Weber, Oren Freifeld*
36. Quantum Motion Segmentation, *Federica Arrigoni, Willi Menapace, Marcel Seelbach Benkner, Elisa Ricci, Vladislav Golyanik*
37. StyleBabel: Artistic Style Tagging and Captioning, *Dan Ruta, Andrew Gilbert, Pranav Aggarwal, Naveen Marri, Ajinkya Kale, Jo Briggs, Chris Speed, Hailin Jin, Baldo Faieta, Alex Filipkowski, Zhe Lin, John Collomosse*
38. VisageSynTalk: Unseen Speaker Video-to-Speech Synthesis via Speech-Visage Feature Selection, *Joanna Hong, Minsu Kim, Yong Man Ro*
39. Semantic Novelty Detection via Relational Reasoning, *Francesco Cappio Borlino, Silvia Bucci, Tatiana Tommasi*
40. Drive&Segment: Unsupervised Semantic Segmentation of Urban Scenes via Cross-Modal Distillation, *Antonin Vobecky, David Hurych, Oriane Siméoni, Spyros Gidaris, Andrei Bursuc, Patrick Pérez, Josef Sivic*
41. POP: Mining POTential Performance of New Fashion Products via Webly Cross-Modal Query Expansion, *Christian Joppi, Geri Skenderi, Marco Cristani*
42. Photo-Realistic Neural Domain Randomization, *Sergey Zakharov, Rareş Ambruş, Vitor Guizilini, Wadim Kehl, Adrien Gaidon*
43. What Matters for 3D Scene Flow Network, *Guangming Wang, Yunzhe Hu, Zhe Liu, Yiyang Zhou, Masayoshi Tomizuka, Wei Zhan, Hesheng Wang*
44. Video Instance Segmentation via Multi-Scale Spatio-Temporal Split Attention Transformer, *Omkar Thawakar, Sanath Narayan, Jiale Cao, Hisham Cholakkal, Rao Muhammad Anwer, Muhammad Haris Khan, Salman Khan, Michael Felsberg, Fahad Shahbaz Khan*
45. Revisiting Batch Norm Initialization, *Jim Davis, Logan Frank*
46. Webly Supervised Concept Expansion for General Purpose Vision Models, *Amita Kamath, Christopher Clark, Tanmay Gupta, Eric Kolve, Derek Hoiem, Aniruddha Kembhavi*
47. Compositional Human-Scene Interaction Synthesis with Semantic Control, *Kaifeng Zhao, Shaofei Wang, Yan Zhang, Thabo Beeler, Siyu Tang*
48. Tailoring Self-Supervision for Supervised Learning, *WonJun Moon, Ji-Hwan Kim, Jae-Pil Heo*
49. Difficulty-Aware Simulator for Open Set Recognition, *WonJun Moon, Junho Park, Hyun Seok Seong, Cheol-Ho Cho, Jae-Pil Heo*
50. SpOT: Spatiotemporal Modeling for 3D Object Tracking, *Colton Stearns, Davis Rempe, Jie Li, Rareş Ambruş, Sergey Zakharov, Vitor Guizilini, Yanchao Yang, Leonidas J. Guibas*
51. VQGAN-CLIP: Open Domain Image Generation and Editing with Natural Language Guidance, *Katherine Crowson, Stella Biderman, Daniel Kornis, Dashiell Stander, Eric Hallahan, Louis Castricato, Edward Raff*
52. Unsupervised High-Fidelity Facial Texture Generation and Reconstruction, *Ron Slossberg, Ibrahim Jubran, Ron Kimmel*
53. DenseHybrid: Hybrid Anomaly Detection for Dense Open-Set Recognition, *Matej Grcić, Petra Bevandić, Siniša Šegvić*
54. Uncertainty-Guided Source-Free Domain Adaptation, *Subhankar Roy, Martin Trapp, Andrea Pilzer, Juho Kannala, Nicu Sebe, Elisa Ricci, Arno Solin*
55. HM: Hybrid Masking for Few-Shot Segmentation, *Seonghyeon Moon, Samuel S. Sohn, Honglu Zhou, Sejong Yoon, Vladimir Pavlovic, Muhammad Haris Khan, Mubbasir Kapadia*
56. Masked Siamese Networks for Label-Efficient Learning, *Mahmoud Assran, Mathilde Caron, Ishan Misra, Piotr Bojanowski, Florian Bordes, Pascal Vincent, Armand Joulin, Michael Rabbat, Nicolas Ballas*
57. FairStyle: Debiasing StyleGAN2 with Style Channel Manipulations, *Cemre Efe Karakas, Alara Dirik, Eylül Yalçinkaya, Pinar Yanardag*
58. Super-Resolution 3D Human Shape from a Single Low-Resolution Image, *Marco Pesavento, Marco Volino, Adrian Hilton*
59. MINER: Multiscale Implicit Neural Representation, *Vishwanath Saragadam, Jasper Tan, Guha Balakrishnan, Richard G. Baraniuk, Ashok Veeraraghavan*

60. Synthesizing Light Field Video from Monocular Video, *Shrisudhan Govindarajan, Prasan Shedligeri, Sarah, Kaushik Mitra*
61. Towards Accurate Open-Set Recognition via Background-Class Regularization, *Wonwoo Cho, Jaegul Choo*
62. Towards Learning Neural Representations from Shadows, *Kushagra Tiwary, Tzofi Klinghoffer, Ramesh Raskar*
63. Augmenting Deep Classifiers with Polynomial Neural Networks, *Grigorios G. Chrysos, Markos Georgopoulos, Jiankang Deng, Jean Kossaifi, Yannis Panagakis, Anima Anandkumar*
64. Minimal Neural Atlas: Parameterizing Complex Surfaces with Minimal Charts and Distortion, *Weng Fei Low, Gim Hee Lee*
65. Video Mask Transfomer for High-Quality Video Instance Segmentation, *Lei Ke, Henghui Ding, Martin Danelljan, Yu-Wing Tai, Chi-Keung Tang, Fisher Yu*
66. Domain Adaptive Hand Keypoint and Pixel Localization in the Wild, *Takehiko Ohkawa, Yu-Jhe Li, Qichen Fu, Ryosuke Furuta, Kris M. Kitani, Yoichi Sato*
67. Meta-Sampler: Almost-Universal yet Task-Oriented Sampling for Point Clouds, *Ta-Ying Cheng, Qingyong Hu, Qian Xie, Niki Trigoni, Andrew Markham*
68. MPIB: An MPI-Based Bokeh Rendering Framework for Realistic Partial Occlusion Effects, *Juewen Peng, Jianming Zhang, Xianrui Luo, Hao Lu, Ke Xian, Zhiguo Cao*
69. Housekeep: Tidying Virtual Households Using Commonsense Reasoning, *Yash Kant, Arun Ramachandran, Sriram Yenamandra, Igor Gilitschenski, Dhruv Batra, Andrew Szot, Harsh Agrawal*
70. LESS: Label-Efficient Semantic Segmentation for LiDAR Point Clouds, *Minghua Liu, Yin Zhou, Charles R. Qi, Boqing Gong, Hao Su, Dragomir Anguelov*
71. Ultra-High-Resolution Unpaired Stain Transformation via Kernelized Instance Normalization, *Ming-Yang Ho, Min-Sheng Wu, Che-Ming Wu*
72. Reliable Visual Question Answering: Abstain Rather Than Answer Incorrectly, *Spencer Whitehead, Suzanne Petryk, Vedaad Shakib, Joseph Gonzalez, Trevor Darrell, Anna Rohrbach, Marcus Rohrbach*
73. A Real World Dataset for Multi-View 3D Reconstruction, *Rakesh Shrestha, Siqi Hu, Minghao Gou, Ziyuan Liu, Ping Tan*
74. CLIP-Actor: Text-Driven Recommendation and Stylization for Animating Human Meshes, *Kim Youwang, Kim Ji-Yeon, Tae-Hyun Oh*
75. NeFSAC: Neurally Filtered Minimal Samples, *Luca Cavalli, Marc Pollefeys, Daniel Barath*
76. DeiT III: Revenge of the ViT, *Hugo Touvron, Matthieu Cord, Hervé Jégou*
77. Map-Free Visual Relocalization: Metric Pose Relative to a Single Image, *Eduardo Arnold, Jamie Wynn, Sara Vicente, Guillermo Garcia-Hernando, Aron Monszpart, Victor Prisacariu, Daniyar Turmukhambetov, Eric Brachmann*
78. Global-Local Motion Transformer for Unsupervised Skeleton-Based Action Learning, *Boeun Kim, Hyung Jin Chang, Jung Ho Kim, Jin Young Choi*
79. The One Where They Reconstructed 3D Humans and Environments in TV Shows, *Georgios Pavlakos, Ethan Weber, Matthew Tancik, Angjoo Kanazawa*
80. EvAC3D: From Event-Based Apparent Contours to 3D Models via Continuous Visual Hulls, *Ziyun Wang, Kenneth Chaney, Kostas Daniilidis*
81. Discovering Deformable Keypoint Pyramids, *Jianing Qian, Anastasios Panagopoulos, Dinesh Jayaraman*
82. Data Invariants to Understand Unsupervised Out-of-Distribution Detection, *Lars Doorenbos, Raphael Sznitman, Pablo Márquez-Neila*
83. Delta Distillation for Efficient Video Processing, *Amirhossein Habibian, Haitam Ben Yahia, Davide Abati, Efstratios Gavves, Fatih Porikli*
84. Completely Self-Supervised Crowd Counting via Distribution Matching, *Deepak Babu Sam, Abhinav Agarwalla, Jimmy Joseph, Vishwanath A. Sindagi, R. Venkatesh Babu, Vishal M. Patel*
85. CoGS: Controllable Generation and Search from Sketch and Style, *Cusuh Ham, Gemma Canet Tarrés, Tu Bui, James Hays, Zhe Lin, John Collomosse*
86. LidarNAS: Unifying and Searching Neural Architectures for 3D Point Clouds, *Chenxi Liu, Zhaoqi Leng, Pei Sun, Shuyang Cheng, Charles R. Qi, Yin Zhou, Mingxing Tan, Dragomir Anguelov*
87. PT4AL: Using Self-Supervised Pretext Tasks for Active Learning, *John Seon Keun Yi, Minseok Seo, Jongchan Park, Dong-Geol Choi*
88. Style-Agnostic Reinforcement Learning, *Juyong Lee, Seokjun Ahn, Jaesik Park*
89. Online Domain Adaptation for Semantic Segmentation in Ever-Changing Conditions, *Theodoros Panagiotakopoulos, Pier Luigi Dovesi, Linus Härenstam-Nielsen, Matteo Poggi*
90. Sim-2-Sim Transfer for Vision-and-Language Navigation in Continuous Environments, *Jacob Krantz, Stefan Lee*
91. Open Vocabulary Object Detection with Pseudo Bounding-Box Labels, *Mingfei Gao, Chen Xing, Juan Carlos Nieves, Junnan Li, Ran Xu, Wenhao Liu, Caiming Xiong*
92. Towards Racially Unbiased Skin Tone Estimation via Scene Disambiguation, *Haiwen Feng, Timo Bolkart, Joachim Tesch, Michael J. Black, Victoria Abrevaya*
93. Combining Internal and External Constraints for Unrolling Shutter in Videos, *Eyal Naor, Itai Antebi, Shai Bagon, Michal Irani*
94. RepMix: Representation Mixing for Robust Attribution of Synthesized Images, *Tu Bui, Ning Yu, John Collomosse*
95. CXR Segmentation by AdaIN-Based Domain Adaptation and Knowledge Distillation, *Yujin Oh, Jong Chul Ye*
96. Diverse Generation from a Single Video Made Possible, *Niv Haim, Ben Feinstein, Niv Granot, Assaf Shocher, Shai Bagon, Tali Dekel, Michal Irani*
97. BodySLAM: Joint Camera Localisation, Mapping, and Human Motion Tracking, *Dorian F. Henning, Tristan Laidlow, Stefan Leutenegger*
98. Differentiable Zooming for Multiple Instance Learning on Whole-Slide Images, *Kevin Thandiackal, Boqi Chen, Pushpak Pati, Guillaume Jaume, Drew F. K. Williamson, Maria Gabrani, Orcun Goksel*
99. Tomography of Turbulence Strength Based on Scintillation Imaging, *Nir Shaul, Yoav Y. Schechner*
100. Pixel-Wise Energy-Biased Abstention Learning for Anomaly Segmentation on Complex Urban Driving Scenes, *Yu Tian, Yuyuan Liu, Guansong Pang, Fengbei Liu, Yuanhong Chen, Gustavo Carneiro*
101. No Token Left Behind: Explainability-Aided Image Classification and Generation, *Roni Paiss, Hila Chefer, Lior Wolf*

Thursday, October 27

Tuesday, October 25	7000	0800	0830	0900	0915	0930	0945	1000	1015	1030	1045	1100	1115	1130	1145	1200	1215	1230	1245	1300	1315	1330		
	Registration (Lobby)																							
		Coffee & Refreshments (Lobby & Hall C)		Exhibits (Hall C): Same on Wednesday and Thursday; see online for details																				
				Demos (Hall C Demo Area), pg. 3																				
				ECCV 2022 At-a-Glance (Main Conference)										Coffee & Refreshments (Lobby & Hall C)					Lunch (Lobby & Hall C)					
Poster Setup (Hall B)		Oral 1.A.1: <i>Detection, Recognition, Classification, & Localization in 2D/3D</i> (Hall D), pg. 3												Mentoring Session (Halls K-M)										
		Oral 1.A.2: <i>Motion and Tracking</i> (Hall E), pg. 3				Poster 1.A (Hall B): pgs. 3-6																		
7000	0800	0830	0900	0915	0930	0945	1000	1015	1030	1045	1100	1115	1130	1145	1200	1215	1230	1245	1300	1315	1330			
Wednesday, October 26	Registration (Lobby)																							
		Coffee & Refreshments (Lobby & Hall C)		Demos (Hall C Demo Area), pg. 11																				
				Oral 2.A.1: <i>Image/Video Synthesis and Generative Models</i> (Hall D), pg. 11										Coffee & Refreshments (Lobby & Hall C)					Lunch (Lobby & Hall C)					
														Poster Setup (Hall B)		Oral 2.A.2: <i>Faces, Bodies, Gestures, and Pose</i> (Hall E), pg. 11				Mentoring Session (Halls K-M)				
	Poster 2.A (Hall B): pgs. 11-14																							
												Industry Track (Hall E): pg. 11												
7000	0800	0830	0900	0915	0930	0945	1000	1015	1030	1045	1100	1115	1130	1145	1200	1215	1230	1245	1300	1315	1330			
Thursday, October 27	Registration (Lobby)																							
		Coffee & Refreshments (Lobby & Hall C)		Demos (Hall C Demo Area), pg. 19																				
				Oral 3.A.1: <i>Stereo and 3D Multiview/Sensors</i> (Hall D), pg. 19										Coffee & Refreshments (Lobby & Hall C)					Lunch (Lobby & Hall C)					
														Poster Setup (Hall B)		Oral 3.A.2: <i>Datasets & Evaluation</i> (Hall E), pg. 19				Poster 3.A (Hall B): pgs. 19-22				
7000	0800	0830	0900	0915	0930	0945	1000	1015	1030	1045	1100	1115	1130	1145	1200	1215	1230	1245	1300	1315	1330			

1345	1400	1415	1430	1445	1500	1515	1530	1545	1600	1615	1630	1645	1700	1715	1730	1745	1800	1830	1900	2000														
y)																																		
r exhibitor list & map																																		
							Coffee & Refreshments (Lobby & Hall C)																											
Poster Swith / Setup (Hall B)							Poster 1.B (Hall B): pgs. 7-10								Opening Remarks & Paper Awards (Halls D & E)																			
Oral 1.B.1: <i>Architecture, Training, and Optimization</i> (Hall D), pg. 7																																		
Oral 1.B.2: <i>Shape From-X and Applications</i> (Hall E), pg. 7																																		
1345	1400	1415	1430	1445	1500	1515	1530	1545	1600	1615	1630	1645	1700	1715	1730	1745	1800	1830	1900	2000														
y)																		Transfer to Reception		Reception (Hanger 11, TLV Port), pg. 18														
							Coffee & Refreshments (Lobby & Hall C)						Keynote Talk: <i>The Science of Ethical Algorithm Design</i> Michael Kearns (Univ. of Pennsylvania) (Halls D & E), pg. 18																					
Poster Swith / Setup (Halls B)							Poster 2.B (Hall B): pgs. 15-18																											
Oral 2.B.1: <i>Scene, Action, and Video Understanding</i> (Hall D), pg. 15																																		
Oral 2.B.2: <i>Low-Level Vision, Segmentation, and Synthesis</i> (Hall E), pg. 15																								Industry Track (Halls K-M): pg. 15										
1345	1400	1415	1430	1445	1500	1515	1530	1545	1600	1615	1630	1645	1700	1715	1730	1745	1800	1830	1900	2000														
y)																		Happy Hour (Lobby & Hall C)																
							Coffee & Refreshments (Lobby & Hall C)						Keynote Talk: <i>Thinking with the Body and the World</i> Barbara Tversky (Stanford Univ. & Columbia Teachers College) (Halls D & E), pg. 26																					
Poster Swith / Setup (Halls B)							Poster 3.B (Hall B): pgs. 23-26																											
Oral 3.B.1: <i>Robots, Vehicles, and Computational Photography</i> (Hall D), pg. 23																																		
Oral 3.B.2: <i>Vision, Text, and Non-Supervised Learning</i> (Hall E), pg. 23																																		
1345	1400	1415	1430	1445	1500	1515	1530	1545	1600	1615	1630	1645	1700	1715	1730	1745	1800	1830	1900	2000														

-
- A full-page sheet of white graph paper with a light gray grid. The grid consists of small squares, approximately 10 units wide by 10 units high. There are no margins or additional markings on the page.

Keynote: Thinking with the Body and the World, *Barbara Tversky (Stanford Univ. & Columbia Teachers College)*

Notes:

[illegible]

Virtual Papers – Oral

Virtual Orals (Online Platform)

These are virtual only papers of accepted ECCV oral papers grouped and sorted by primary topic.

Format (5 min. video presentation; online poster)

3D from a Single Image and Shape-from-X

- Organic Priors in Non-rigid Structure from Motion, *Suryansh Kumar, Luc Van Gool*

3D Shape Modeling and Processing

- FBNet: Feedback Network for Point Cloud Completion, *Xuejun Yan, Hongyu Yan, Jingjing Wang, Hang Du, Zhihong Wu, Di Xie, Shiliang Pu, Li Lu*

Action and Behavior Recognition

- CMD: Self-Supervised 3D Action Representation Learning with Cross-Modal Mutual Distillation, *Yunyao Mao, Wengang Zhou, Zhenbo Lu, Jiajun Deng, Houqiang Li*
- Expanding Language-Image Pretrained Models for General Video Recognition, *Bolin Ni, Houwen Peng, Minghao Chen, Songyang Zhang, Gaofeng Meng, Jianlong Fu, Shiming Xiang, Haibin Ling*
- Delving into Details: Synopsis-to-Detail Networks for Video Recognition, *Shuxian Liang, Xu Shen, Jianqiang Huang, Xian-Sheng Hua*

Adversarial Learning

- Frequency Domain Model Augmentation for Adversarial Attack, *Yuyang Long, Qilong Zhang, Boheng Zeng, Lianli Gao, Xianglong Liu, Jian Zhang, Jingkuan Song*

Biometrics

- UIA-ViT: Unsupervised Inconsistency-Aware Method Based on Vision Transformer for Face Forgery Detection, *Wanyi Zhuang, Qi Chu, Zhentao Tan, Qiankun Liu, Haojie Yuan, Changtao Miao, Zixiang Luo, Nenghai Yu*

Body Gestures and Pose

- D&D: Learning Human Dynamics from Dynamic Camera, *Jiefeng Li, Siyuan Bian, Chao Xu, Gang Liu, Gang Yu, Cewu Lu*
- SimCC: A Simple Coordinate Classification Perspective for Human Pose Estimation, *Yanjie Li, Sen Yang, Peidong Liu, Shoukui Zhang, Yunxiao Wang, Zhicheng Wang, Wankou Yang, Shu-Tao Xia*
- Deep Radial Embedding for Visual Sequence Learning, *Yuecong Min, Peiqi Jiao, Yanan Li, Xiaotao Wang, Lei Lei, Xiujuan Chai, Xilin Chen*

Computational Photography

- Estimating Spatially-Varying Lighting in Urban Scenes with Disentangled Representation, *Jiajun Tang, Yongjie Zhu, Haoyu Wang, Jun Hoong Chan, Si Li, Boxin Shi*
- CT²: Colorization Transformer via Color Tokens, *Shuchen Weng, Jimeng Sun, Yu Li, Si Li, Boxin Shi*
- Bringing Rolling Shutter Images Alive with Dual Reversed Distortion, *Zhihang Zhong, Mingdeng Cao, Xiao Sun, Zhirong Wu, Zhongyi Zhou, Yinqiang Zheng, Stephen Lin, Imari Sato*
- DCCF: Deep Comprehensible Color Filter Learning Framework for High-Resolution Image Harmonization, *Ben Xue, Shenghui Ran, Quan Chen, Rongfei Jia, Binqiang Zhao, Xing Tang*

Datasets and Evaluation

- 3D CoMPaT: Composition of Materials on Parts of 3D Things, *Yuchen Li, Ujjwal Upadhyay, Habib Slim, Tezvesh Varshney,*

Ahmed Abdelreheem, Arpit Prajapati, Suhail Pothigara, Peter Wonka, Mohamed Elhoseiny

- OOD-CV: A Benchmark for Robustness to Out-of-Distribution Shifts of Individual Nuisances in Natural Images, *Bingchen Zhao, Shaozuo Yu, Wufei Ma, Mingxin Yu, Shenxiao Mei, Angtian Wang, Ju He, Alan Yuille, Adam Kortylewski*

Detection and Localization in 2D and/or 3D

- Graph R-CNN: Towards Accurate 3D Object Detection with Semantic-Decorated Local Graph, *Honghui Yang, Zili Liu, Xiaopei Wu, Wenxiao Wang, Wei Qian, Xiaofei He, Deng Cai*
- Open-Vocabulary DETR with Conditional Matching, *Yuhang Zang, Wei Li, Kaiyang Zhou, Chen Huang, Chen Change Loy*
- Adversarially-Aware Robust Object Detector, *Ziyi Dong, Pengxu Wei, Liang Lin*
- Cross-Modality Knowledge Distillation Network for Monocular 3D Object Detection, *Yu Hong, Hang Dai, Yong Ding*
- 3D Object Detection with a Self-Supervised Lidar Scene Flow Backbone, *Emeç Erçelik, Ekim Yurtsever, Mingyu Liu, Zhijie Yang, Hanzhen Zhang, Pinar Topçam, Maximilian Listl, Yılmaz Kaan Çaylı, Alois Knoll*
- ObjectBox: From Centers to Boxes for Anchor-Free Object Detection, *Mohsen Zand, Ali Etemad, Michael Greenspan*

Faces

- On Mitigating Hard Clusters for Face Clustering, *Yingjie Chen, Huasong Zhong, Chong Chen, Chen Shen, Jianqiang Huang, Tao Wang, Yun Liang, Qianru Sun*

Image and Video Retrieval

- Domain Adaptive Person Search, *Junjie Li, Yichao Yan, Guanshuo Wang, Fufu Yu, Qiong Jia, Shouhong Ding*

Image and Video Synthesis

- Contrastive Monotonic Pixel-Level Modulation, *Kun Lu, Rongpeng Li, Honggang Zhang*
- Discovering Transferable Forensic Features for CNN-Generated Images Detection, *Keshigeyan Chandrasegaran, Ngoc-Trung Tran, Alexander Binder, Ngai-Man Cheung*
- Auto-Regressive Image Synthesis with Integrated Quantization, *Fangneng Zhan, Yingchen Yu, Rongliang Wu, Jiahui Zhang, Kaiwen Cui, Changgong Zhang, Shijian Lu*
- CCPL: Contrastive Coherence Preserving Loss for Versatile Style Transfer, *Zijie Wu, Zhen Zhu, Junping Du, Xiang Bai*
- NeuMesh: Learning Disentangled Neural Mesh-Based Implicit Field for Geometry and Texture Editing, *Bangbang Yang, Chong Bao, Junyi Zeng, Hujun Bao, Yinda Zhang, Zhaopeng Cui, Guofeng Zhang*
- NÜWA: Visual Synthesis Pre-training for Neural visual World Generation, *Chenfei Wu, Jian Liang, Lei Ji, Fan Yang, Yuejian Fang, Daxin Jiang, Nan Duan*
- Neural Radiance Transfer Fields for Relightable Novel-View Synthesis with Global Illumination, *Linjie Lyu, Ayush Tewari, Thomas Leimkühler, Marc Habermann, Christian Theobalt*

Low-Level and Physics-Based Vision

- VQFR: Blind Face Restoration with Vector-Quantized Dictionary and Parallel Decoder, *Yuchao Gu, Xintao Wang, Liangbin Xie, Chao Dong, Gen Li, Ying Shan, Ming-Ming Cheng*
- Adaptive Patch Exiting for Scalable Single Image Super-Resolution, *Shizun Wang, Jiaming Liu, Kaixin Chen, Xiaoqi Li, Ming Lu, Yandong Guo*
- Event-Guided Deblurring of Unknown Exposure Time Videos, *Taewoo Kim, Jeongmin Lee, Lin Wang, Kuk-Jin Yoon*

- Restore Globally, Refine Locally: A Mask-Guided Scheme to Accelerate Super-Resolution Networks, *Xiaotao Hu, Jun Xu, Shuhang Gu, Ming-Ming Cheng, Li Liu*
- Modeling Mask Uncertainty in Hyperspectral Image Reconstruction, *Jiamian Wang, Yulun Zhang, Xin Yuan, Ziyi Meng, Zhiqiang Tao*
- Perceiving and Modeling Density for Image Dehazing, *Tian Ye, Yunchen Zhang, Mingchao Jiang, Liang Chen, Yun Liu, Sixiang Chen, Erkang Chen*
- Stripformer: Strip Transformer for Fast Image Deblurring, *Fu-Jen Tsai, Yan-Tsung Peng, Yen-Yu Lin, Chung-Chi Tsai, Chia-Wen Lin*
- Deep Fourier-Based Exposure Correction Network with Spatial-Frequency Interaction, *Jie Huang, Yajing Liu, Feng Zhao, Keyu Yan, Jinghao Zhang, Yukun Huang, Man Zhou, Zhiwei Xiong*
- KXNet: A Model-Driven Deep Neural Network for Blind Super-Resolution, *Jiahong Fu, Hong Wang, Qi Xie, Qian Zhao, Deyu Meng, Zongben Xu*
- RealFlow: EM-Based Realistic Optical Flow Dataset Generation from Videos, *Yunhui Han, Kunming Luo, Ao Luo, Jiangyu Liu, Haoqiang Fan, Guiming Luo, Shuaicheng Liu*
- RRSR: Reciprocal Reference-Based Image Super-Resolution with Progressive Feature Alignment and Selection, *Lin Zhang, Xin Li, Dongliang He, Fu Li, Yili Wang, Zhaoxiang Zhang*

Medical, Biological, and Cell Microscopy

- **PointScatter**: Point Set Representation for Tubular Structure Extraction, *Dong Wang, Zhao Zhang, Ziwei Zhao, Yuhang Liu, Yihong Chen, Liwei Wang*

Motion and Tracking

- Towards Grand Unification of Object Tracking, *Bin Yan, Yi Jiang, Peize Sun, Dong Wang, Zehuan Yuan, Ping Luo, Huchuan Lu*
- Tracking Objects as Pixel-Wise Distributions, *Zelin Zhao, Ze Wu, Yueqing Zhuang, Boxun Li, Jiaya Jia*

Recognition and Classification

- Recurrent Bilinear Optimization for Binary Neural Networks, *Sheng Xu, Yanjing Li, Tiancheng Wang, Teli Ma, Baochang Zhang, Peng Gao, Yu Qiao, Jinhu Lü, Guodong Guo*
- Towards Calibrated Hyper-Sphere Representation via Distribution Overlap Coefficient for Long-Tailed Learning, *Hualiang Wang, Siming Fu, Xiaoxuan He, Hangxiang Fang, Zuozhu Liu, Haoji Hu*
- AutoMix: Unveiling the Power of Mixup for Stronger Classifiers, *Zicheng Liu, Siyuan Li, Di Wu, Zihan Liu, Zhiyuan Chen, Lirong Wu, Stan Z. Li*

Representation Learning

- Contrastive Deep Supervision, *Lin Feng Zhang, Xin Chen, Junbo Zhang, Runpei Dong, Kaisheng Ma*
- Balancing Stability and Plasticity through Advanced Null Space in Continual Learning, *Yajing Kong, Liu Liu, Zhen Wang, Dacheng Tao*
- DisCo: Remedying Self-Supervised Learning on Lightweight Models with Distilled Contrastive Learning, *Yuting Gao, Jia-Xin Zhuang, Shaohui Lin, Hao Cheng, Xing Sun, Ke Li, Chunhua Shen*
- Identifying Hard Noise in Long-Tailed Sample Distribution, *Xuanyu Yi, Kaihua Tang, Xian-Sheng Hua, Joo-Hwee Lim, Hanwang Zhang*

Scene Analysis and Understanding

- TO-Scene: A Large-Scale Dataset for Understanding 3D Tabletop Scenes, *Mutian Xu, Pei Chen, Haolin Liu, Xiaoguang Han*
- Fine-Grained Scene Graph Generation with Data Transfer, *Ao Zhang, Yuan Yao, Qianyu Chen, Wei Ji, Zhiyuan Liu, Maosong Sun, Tat-Seng Chua*

Scene Text and Document Understanding

- Language Matters: A Weakly Supervised Vision-Language Pre-training Approach for Scene Text Detection and Spotting, *Chuhui Xue, Wenqing Zhang, Yu Hao, Shijian Lu, Philip H. S. Torr, Song Bai*
- Toward Understanding WordArt: Corner-Guided Transformer for Scene Text Recognition, *Xudong Xie, Ling Fu, Zhifei Zhang, Zhaowen Wang, Xiang Bai*

Segmentation, Grouping, and Shape

- SeqFormer: Sequential Transformer for Video Instance Segmentation, *Junfeng Wu, Yi Jiang, Song Bai, Wenqing Zhang, Xiang Bai*
- In Defense of Online Models for Video Instance Segmentation, *Junfeng Wu, Qihao Liu, Yi Jiang, Song Bai, Alan Yuille, Xiang Bai*
- Extract Free Dense Labels from CLIP, *Chong Zhou, Chen Change Loy, Bo Dai*
- TransFGU: A Top-down Approach to Fine-Grained Unsupervised Semantic Segmentation, *Zhaoyuan Yin, Pichao Wang, Fan Wang, Xianzhe Xu, Hanling Zhang, Hao Li, Rong Jin*
- BATMAN: Bilateral Attention Transformer in Motion-Appearance Neighboring Space for Video Object Segmentation, *Ye Yu, Jialin Yuan, Gaurav Mittal, Li Fuxin, Mei Chen*
- Learning Topological Interactions for Multi-Class Medical Image Segmentation, *Saumya Gupta, Xiaoling Hu, James Kaan, Michael Jin, Mutshipay Mpoy, Katherine Chung, Gagandeep Singh, Mary Saltz, Tahsin Kurc, Joel Saltz, Apostolos Tassiopoulos, Prateek Prasanna, Chao Chen*
- Unsupervised Segmentation in Real-World Images via Spelke Object Inference, *Honglin Chen, Rahul Venkatesh, Yoni Friedman, Jiajun Wu, Joshua B. Tenenbaum, Daniel L. K. Yamins, Daniel M. Bear*

Semi / Weak / Self / Unsupervised Learning

- Open-Set Semi-Supervised Object Detection, *Yen-Cheng Liu, Chih-Yao Ma, Xiaoliang Dai, Junjiao Tian, Peter Vajda, Zijian He, Zsolt Kira*
- Pointly-Supervised Panoptic Segmentation, *Junsong Fan, Zhaoxiang Zhang, Tieniu Tan*
- Synergistic Self-Supervised and Quantization Learning, *Yun-Hao Cao, Peiqin Sun, Yechang Huang, Jianxin Wu, Shuchang Zhou*
- Semi-Supervised Object Detection via Virtual Category Learning, *Changrui Chen, Kurt Debattista, Jungong Han*
- Improving Self-Supervised Lightweight Model Learning via Hard-Aware Metric Distillation, *Hao Liu, Mang Ye*
- Towards Realistic Semi-Supervised Learning, *Mamshad Nayeem Rizve, Navid Kardan, Mubarak Shah*

Stereo, 3D from Multiview and Other Sensors

- PCW-Net: Pyramid Combination and Warping Cost Volume for Stereo Matching, *Zhelun Shen, Yuchao Dai, Xibin Song, Zhibo Rao, Dingfu Zhou, Liangjun Zhang*
- Affine Correspondences between Multi-Camera Systems for 6DOF Relative Pose Estimation, *Banglei Guan, Ji Zhao*

- [illegible]

- [illegible]

[illegible]

- [illegible]

[illegible]

-

- [illegible]

[illegible]

Virtual Papers – Poster

Virtual Posters (Online Platform)

These are virtual only papers of accepted ECCV poster papers grouped and sorted by primary topic.

Format (5 min. video presentation; online poster)

3D from a Single Image and Shape-from-X

- Learning-Based Point Cloud Registration for 6D Object Pose Estimation in the Real World, *Zheng Dang, Lizhou Wang, Yu Guo, Mathieu Salzmann*
- An End-to-End Transformer Model for Crowd Localization, *Dingkang Liang, Wei Xu, Xiang Bai*
- Few-Shot Single-View 3D Reconstruction with Memory Prior Contrastive Network, *Zhen Xing, Yijiang Chen, Zhixin Ling, Xiangdong Zhou, Yu Xiang*
- DID-M3D: Decoupling Instance Depth for Monocular 3D Object Detection, *Liang Peng, Xiaopei Wu, Zheng Yang, Haifeng Liu, Deng Cai*
- Adaptive Co-Teaching for Unsupervised Monocular Depth Estimation, *Weisong Ren, Lijun Wang, Yongri Piao, Miao Zhang, Huchuan Lu, Ting Liu*
- Lidar Point Cloud Guided Monocular 3D Object Detection, *Liang Peng, Fei Liu, Zhengxu Yu, Senbo Yan, Dan Deng, Zheng Yang, Haifeng Liu, Deng Cai*
- 3D Human Pose Estimation Using Möbius Graph Convolutional Networks, *Niloofer Azizi, Horst Possegger, Emanuele Rodolà, Horst Bischof*
- Learning to Train a Point Cloud Reconstruction Network without Matching, *Tianxin Huang, Xuemeng Yang, Jiangning Zhang, Jinhao Cui, Hao Zou, Jun Chen, Xiangrui Zhao, Yong Liu*
- PanoFormer: Panorama Transformer for Indoor 360° Depth Estimation, *Zhijie Shen, Chunyu Lin, Kang Liao, Lang Nie, Zishuo Zheng, Yao Zhao*
- Towards Comprehensive Representation Enhancement in Semantics-Guided Self-Supervised Monocular Depth Estimation, *Jingyuan Ma, Xiangyu Lei, Nan Liu, Xian Zhao, Shiliang Pu*
- AvatarCap: Animatable Avatar Conditioned Monocular Human Volumetric Capture, *Zhe Li, Zerong Zheng, Hongwen Zhang, Chaonan Ji, Yebin Liu*
- GeoRefine: Self-Supervised Online Depth Refinement for Accurate Dense Mapping, *Pan Ji, Qingan Yan, Yuxin Ma, Yi Xu*
- Multi-modal Masked Pre-training for Monocular Panoramic Depth Completion, *Zhiqiang Yan, Xiang Li, Kun Wang, Zhenyu Zhang, Jun Li, Jian Yang*
- GitNet: Geometric Prior-Based Transformation for Birds-Eye-View Segmentation, *Shi Gong, Xiaoqing Ye, Xiao Tan, Jingdong Wang, Errui Ding, Yu Zhou, Xiang Bai*
- Learning Visibility for Robust Dense Human Body Estimation, *Chun-Han Yao, Jimei Yang, Duygu Ceylan, Yi Zhou, Yang Zhou, Ming-Hsuan Yang*
- Towards High-Fidelity Single-View Holistic Reconstruction of Indoor Scenes, *Haolin Liu, Yujian Zheng, Guanying Chen, Shuguang Cui, Xiaoguang Han*
- SketchSampler: Sketch-Based 3D Reconstruction via View-Dependent Depth Sampling, *Chenjian Gao, Qian Yu, Lu Sheng, Yi-Zhe Song, Dong Xu*
- LocalBins: Improving Depth Estimation by Learning Local Distributions, *Shariq Farooq Bhat, Ibraheem Alhashim, Peter Wonka*
- Semi-Supervised Single-View 3D Reconstruction via Prototype Shape Priors, *Zhen Xing, Hengduo Li, Zuxuan Wu, Yu-Gang Jiang*
- SC-wLS: Towards Interpretable Feed-Forward Camera Re-localization, *Xin Wu, Hao Zhao, Shunkai Li, Yingdian Cao, Hongbin Zha*
- FloatingFusion: Depth from ToF and Image-Stabilized Stereo Cameras, *Andreas Meuleman, Hakyeon Kim, James Tompkin, Min H. Kim*
- DELTAR: Depth Estimation from a Light-Weight ToF Sensor and RGB Image, *Yijin Li, Xinyang Liu, Wenqi Dong, Han Zhou, Hujun Bao, Guofeng Zhang, Yinda Zhang, Zhaopeng Cui*
- 3D Room Layout Estimation from a Cubemap of Panorama Image via Deep Manhattan Hough Transform, *Yining Zhao, Chao Wen, Zhou Xue, Yue Gao*
- RBP-Pose: Residual Bounding Box Projection for Category-Level Pose Estimation, *Ruida Zhang, Yan Di, Zhiqiang Lou, Fabian Manhardt, Federico Tombari, Xiangyang Ji*
- Self-Distilled Feature Aggregation for Self-Supervised Monocular Depth Estimation, *Zhengming Zhou, Qiulei Dong*
- MHR-Net: Multiple-Hypothesis Reconstruction of Non-rigid Shapes from 2D Views, *Haitian Zeng, Xin Yu, Jiaxu Miao, Yi Yang*
- Resolution-Free Point Cloud Sampling Network with Data Distillation, *Tianxin Huang, Jiangning Zhang, Jun Chen, Yuang Liu, Yong Liu*
- Self-Calibrating Photometric Stereo by Neural Inverse Rendering, *Junxuan Li, Hongdong Li*
- Object Level Depth Reconstruction for Category Level 6D Object Pose Estimation from Monocular RGB Image, *Zhaoxin Fan, Zhenbo Song, Jian Xu, Zhicheng Wang, Kejian Wu, Hongyan Liu, Jun He*
- Uncertainty Quantification in Depth Estimation via Constrained Ordinal Regression, *Dongting Hu, Lihua Peng, Tingjin Chu, Xiaoxing Zhang, Yinian Mao, Howard Bondell, Mingming Gong*
- CostDCNet: Cost Volume Based Depth Completion for a Single RGB-D Image, *Jaewon Kam, Jungeon Kim, Soongjin Kim, Jaesik Park, Seungyong Lee*
- 3D Siamese Transformer Network for Single Object Tracking on Point Clouds, *Le Hui, Lingpeng Wang, Linghua Tang, Kaihao Lan, Jin Xie, Jian Yang*
- Object Wake-Up: 3D Object Rigging from a Single Image, *Ji Yang, Xinxin Zuo, Sen Wang, Zhenbo Yu, Xingyu Li, Bingbing Ni, Minglun Gong, Li Cheng*
- IntegratedPIFu: Integrated Pixel Aligned Implicit Function for Single-View Human Reconstruction, *Kennard Yanting Chan, Guosheng Lin, Haiyu Zhao, Weisi Lin*
- Neural Light Field Estimation for Street Scenes with Differentiable Virtual Object Insertion, *Zian Wang, Wenzheng Chen, David Acuna, Jan Kautz, Sanja Fidler*
- Perspective Phase Angle Model for Polarimetric 3D Reconstruction, *Guangcheng Chen, Li He, Yisheng Guan, Hong Zhang*
- Camera Auto-Calibration from the Steiner Conic of the Fundamental Matrix, *Yu Liu, Hui Zhang*

3D Shape Modeling and Processing

- CATRE: Iterative Point Clouds Alignment for Category-Level Object Pose Refinement, *Xingyu Liu, Gu Wang, Yi Li, Xiangyang Ji*
- Optimization over Disentangled Encoding: Unsupervised Cross-Domain Point Cloud Completion via Occlusion Factor Manipulation, *Jingyu Gong, Fengqi Liu, Jiachen Xu, Min Wang,*

Xin Tan, Zhizhong Zhang, Ran Yi, Haichuan Song, Yuan Xie, Lizhuang Ma

- Unsupervised Learning of 3D Semantic Keypoints with Mutual Reconstruction, *Haocheng Yuan, Chen Zhao, Shichao Fan, Jiaxi Jiang, Jiaqi Yang*
- Masked Discrimination for Self-Supervised Learning on Point Clouds, *Haotian Liu, Mu Cai, Yong Jae Lee*
- Efficient Point Cloud Analysis Using Hilbert Curve, *Wanli Chen, Xinge Zhu, Guojin Chen, Bei Yu*
- MeshMAE: Masked Autoencoders for 3D Mesh Data Analysis, *Yaqian Liang, Shanshan Zhao, Baosheng Yu, Jing Zhang, Fazhi He*
- Unsupervised Deep Multi-Shape Matching, *Dongliang Cao, Florian Bernard*
- PRIF: Primary Ray-Based Implicit Function, *Brandon Y. Feng, Yinda Zhang, Danhang Tang, Ruofei Du, Amitabh Varshney*
- Point Cloud Domain Adaptation via Masked Local 3D Structure Prediction, *Hanxue Liang, Hehe Fan, Zhiwen Fan, Yi Wang, Tianlong Chen, Yu Cheng, Zhangyang Wang*
- Learning Implicit Templates for Point-Based Clothed Human Modeling, *Siyu Lin, Hongwen Zhang, Zerong Zheng, Ruizhi Shao, Yebin Liu*
- Exploring the Devil in Graph Spectral Domain for 3D Point Cloud Attacks, *Qianjiang Hu, Daizong Liu, Wei Hu*
- Structure-Aware Editable Morphable Model for 3D Facial Detail Animation and Manipulation, *Jingwang Ling, Zhibo Wang, Ming Lu, Quan Wang, Chen Qian, Feng Xu*
- MoFaNeRF: Morphable Facial Neural Radiance Field, *Yiyu Zhuang, Hao Zhu, Xusen Sun, Xun Cao*
- PointInst3D: Segmenting 3D Instances by Points, *Tong He, Wei Yin, Chunhua Shen, Anton van den Hengel*
- Cross-Modal 3D Shape Generation and Manipulation, *Zezhou Cheng, Menglei Chai, Jian Ren, Hsin-Ying Lee, Kyle Olszewski, Zeng Huang, Subhransu Maji, Sergey Tulyakov*
- Latent Partition Implicit with Surface Codes for 3D Representation, *Chao Chen, Yu-Shen Liu, Zhizhong Han*
- PD-Flow: A Point Cloud Denoising Framework with Normalizing Flows, *Aihua Mao, Zihui Du, Yu-Hui Wen, Jun Xuan, Yong-Jin Liu*
- SeedFormer: Patch Seeds Based Point Cloud Completion with Upsample Transformer, *Haoran Zhou, Yun Cao, Wenqing Chu, Junwei Zhu, Tong Lu, Ying Tai, Chengjie Wang*
- A Repulsive Force Unit for Garment Collision Handling in Neural Networks, *Qingyang Tan, Yi Zhou, Tuanfeng Wang, Duygu Ceylan, Xin Sun, Dinesh Manocha*
- SPE-Net: Boosting Point Cloud Analysis via Rotation Robustness Enhancement, *Zhaofan Qiu, Yehao Li, Yu Wang, Yingwei Pan, Ting Yao, Tao Mei*
- The Shape Part Slot Machine: Contact-Based Reasoning for Generating 3D Shapes from Parts, *Kai Wang, Paul Guerrero, Vladimir G. Kim, Siddhartha Chaudhuri, Minhyuk Sung, Daniel Ritchie*

Action and Behavior Recognition

- Spatiotemporal Self-Attention Modeling with Temporal Patch Shift for Action Recognition, *Wangmeng Xiang, Chao Li, Biao Wang, Xihan Wei, Xian-Sheng Hua, Lei Zhang*
- S2N: Suppression-Strengthen Network for Event-Based Recognition under Variant Illuminations, *Zengyu Wan, Yang Wang, Ganchao Tan, Yang Cao, Zheng-Jun Zha*

- Contrastive Positive Mining for Unsupervised 3D Action Representation Learning, *Haoyuan Zhang, Yonghong Hou, Wenjing Zhang, Wanqing Li*
- Uncertainty-Based Spatial-Temporal Attention for Online Action Detection, *Hongji Guo, Zhou Ren, Yi Wu, Gang Hua, Qiang Ji*
- Iwin: Human-Object Interaction Detection via Transformer with Irregular Windows, *Danyang Tu, Xiongkuo Min, Huiyu Duan, Guodong Guo, Guangtao Zhai, Wei Shen*
- Rethinking Zero-Shot Action Recognition: Learning from Latent Atomic Actions, *Yijun Qian, Lijun Yu, Wenhe Liu, Alexander G. Hauptmann*
- Mining Cross-Person Cues for Body-Part Interactiveness Learning in HOI Detection, *Xiaoqian Wu, Yong-Lu Li, Xinpeng Liu, Junyi Zhang, Yuzhe Wu, Cewu Lu*
- Collaborating Domain-Shared and Target-Specific Feature Clustering for Cross-Domain 3D Action Recognition, *Qinying Liu, Zilei Wang*
- Learning Spatial-Preserved Skeleton Representations for Few-Shot Action Recognition, *Ning Ma, Hongyi Zhang, Xuhui Li, Sheng Zhou, Zhen Zhang, Jun Wen, Haifeng Li, Jingjun Gu, Jiajun Bu*
- Dual-Evidential Learning for Weakly-Supervised Temporal Action Localization, *Mengyuan Chen, Junyu Gao, Shicai Yang, Changsheng Xu*
- AdaFocusV3: On Unified Spatial-Temporal Dynamic Video Recognition, *Yulin Wang, Yang Yue, Xinhong Xu, Ali Hassani, Victor Kulikov, Nikita Orlov, Shiji Song, Humphrey Shi, Gao Huang*
- Panoramic Human Activity Recognition, *Ruize Han, Haomin Yan, Jiacheng Li, Songmiao Wang, Wei Feng, Song Wang*
- A Generalized & Robust Framework for Timestamp Supervision in Temporal Action Segmentation, *Rahul Rahaman, Dipika Singhanian, Alexandre Thiery, Angela Yao*
- Few-Shot Action Recognition with Hierarchical Matching and Contrastive Learning, *Sipeng Zheng, Shizhe Chen, Qin Jin*
- Scale-Aware Spatio-Temporal Relation Learning for Video Anomaly Detection, *Guoqiu Li, Guanxiong Cai, Xingyu Zeng, Rui Zhao*
- Compound Prototype Matching for Few-Shot Action Recognition, *Yifei Huang, Lijin Yang, Yoichi Sato*
- Dynamic Spatio-Temporal Specialization Learning for Fine-Grained Action Recognition, *Tianjiao Li, Lin Geng Foo, Qihong Ke, Hossein Rahmani, Anran Wang, Jinghua Wang, Jun Liu*
- Dynamic Local Aggregation Network with Adaptive Clusterer for Anomaly Detection, *Zhiwei Yang, Peng Wu, Jing Liu, Xiaotao Liu*
- Action Quality Assessment with Temporal Parsing Transformer, *Yang Bai, Desen Zhou, Songyang Zhang, Jian Wang, Errui Ding, Yu Guan, Yang Long, Jingdong Wang*
- Pairwise Contrastive Learning Network for Action Quality Assessment, *Mingzhe Li, Hong-Bo Zhang, Qing Lei, Zongwen Fan, Jinghua Liu, Ji-Xiang Du*
- ActionFormer: Localizing Moments of Actions with Transformers, *Chen-Lin Zhang, Jianxin Wu, Yin Li*
- SocialVAE: Human Trajectory Prediction Using Timewise Latents, *Pei Xu, Jean-Bernard Hayet, Ioannis Karamouzas*

Adversarial Learning

- Shape Matters: Deformable Patch Attack, *Zhaoyu Chen, Bo Li, Shuang Wu, Jianghe Xu, Shouhong Ding, Wenqiang Zhang*

- Prior-Guided Adversarial Initialization for Fast Adversarial Training, *Xiaojun Jia, Yong Zhang, Xingxing Wei, Baoyuan Wu, Ke Ma, Jue Wang, Xiaochun Cao*
 - Enhanced Accuracy and Robustness via Multi-Teacher Adversarial Distillation, *Shiji Zhao, Jie Yu, Zhenlong Sun, Bo Zhang, Xingxing Wei*
 - A Large-Scale Multiple-Objective Method for Black-Box Attack against Object Detection, *Siyuan Liang, Longkang Li, Yanbo Fan, Xiaojun Jia, Jingzhi Li, Baoyuan Wu, Xiaochun Cao*
 - A Spectral View of Randomized Smoothing under Common Corruptions: Benchmarking and Improving Certified Robustness, *Jiachen Sun, Akshay Mehra, Bhavya Kailkhura, Pin-Yu Chen, Dan Hendrycks, Jihun Hamm, Z. Morley Mao*
 - Improving Adversarial Robustness of 3D Point Cloud Classification Models, *Guanlin Li, Guowen Xu, Han Qiu, Ruan He, Jiwei Li, Tianwei Zhang*
 - Learning Extremely Lightweight and Robust Model with Differentiable Constraints on Sparsity and Condition Number, *Xian Wei, Yangyu Xu, Yanhui Huang, Hairong Lv, Hai Lan, Mingsong Chen, Xuan Tang*
 - RIBAC: Towards Robust and Imperceptible Backdoor Attack against Compact DNN, *Huy Phan, Cong Shi, Yi Xie, Tianfang Zhang, Zhuohang Li, Tianming Zhao, Jian Liu, Yan Wang, Yingying Chen, Bo Yuan*
 - Boosting Transferability of Targeted Adversarial Examples via Hierarchical Generative Networks, *Xiao Yang, Yinpeng Dong, Tianyu Pang, Hang Su, Jun Zhu*
 - Adaptive Image Transformations for Transfer-Based Adversarial Attack, *Zheng Yuan, Jie Zhang, Shiguang Shan*
 - AdvDO: Realistic Adversarial Attacks for Trajectory Prediction, *Yulong Cao, Chaowei Xiao, Anima Anandkumar, Danfei Xu, Marco Pavone*
 - One Size Does NOT Fit All: Data-Adaptive Adversarial Training, *Shuo Yang, Chang Xu*
 - UniCR: Universally Approximated Certified Robustness via Randomized Smoothing, *Hanbin Hong, Binghui Wang, Yuan Hong*
 - Hardly Perceptible Trojan Attack against Neural Networks with Bit Flips, *Jiawang Bai, Kuofeng Gao, Dihong Gong, Shu-Tao Xia, Zhifeng Li, Wei Liu*
 - Robust Network Architecture Search via Feature Distortion Restraining, *Yaguan Qian, Shenghui Huang, Bin Wang, Xiang Ling, Xiaohui Guan, Zhaoquan Gu, Shaoning Zeng, Wujie Zhou, Haijiang Wang*
 - SecretGen: Privacy Recovery on Pre-trained Models via Distribution Discrimination, *Zhuowen Yuan, Fan Wu, Yunhui Long, Chaowei Xiao, Bo Li*
 - Triangle Attack: A Query-Efficient Decision-Based Adversarial Attack, *Xiaosen Wang, Zeliang Zhang, Kangheng Tong, Dihong Gong, Kun He, Zhifeng Li, Wei Liu*
 - Data-Free Backdoor Removal Based on Channel Lipschitzness, *Runkai Zheng, Rongjun Tang, Jianze Li, Li Liu*
 - Black-Box Dissector: Towards Erasing-Based Hard-Label Model Stealing Attack, *Yixu Wang, Jie Li, Hong Liu, Yan Wang, Yongjian Wu, Feiyue Huang, Rongrong Ji*
 - Learning Energy-Based Models with Adversarial Training, *Xuwang Yin, Shiyang Li, Gustavo K. Rohde*
 - Revisiting Outer Optimization in Adversarial Training, *Ali Dabouei, Fariborz Taherkhani, Sobhan Soleymani, Nasser M. Nasrabadi*
 - Zero-Shot Attribute Attacks on Fine-Grained Recognition Models, *Nasim Shafiee, Ehsan Elhamifar*
 - Towards Effective and Robust Neural Trojan Defenses via Input Filtering, *Kien Do, Haripriya Harikumar, Hung Le, Dung Nguyen, Truyen Tran, Santu Rana, Dang Nguyen, Willy Susilo, Svetha Venkatesh*
- ### Biometrics
- Generative Domain Adaptation for Face Anti-Spoofing, *Qianyu Zhou, Ke-Yue Zhang, Taiping Yao, Ran Yi, Kekai Sheng, Shouhong Ding, Lizhuang Ma*
 - MetaGait: Learning to Learn an Omni Sample Adaptive Representation for Gait Recognition, *Huanzhang Dou, Pengyi Zhang, Wei Su, Yunlong Yu, Xi Li*
 - GaitEdge: Beyond Plain End-to-End Gait Recognition for Better Practicality, *Junhao Liang, Chao Fan, Saihui Hou, Chuanfu Shen, Yongzhen Huang, Shiqi Yu*
 - Effective Presentation Attack Detection Driven by Face Related Task, *Wentian Zhang, Haozhe Liu, Feng Liu, Raghavendra Ramachandra, Christoph Busch*
- ### Body Gestures and Pose
- P-STMO: Pre-trained Spatial Temporal Many-to-One Model for 3D Human Pose Estimation, *Wenkang Shan, Zhenhua Liu, Xinfeng Zhang, Shanshe Wang, Siwei Ma, Wen Gao*
 - Explicit Occlusion Reasoning for Multi-Person 3D Human Pose Estimation, *Qihao Liu, Yi Zhang, Song Bai, Alan Yuille*
 - C3P: Cross-Domain Pose Prior Propagation for Weakly Supervised 3D Human Pose Estimation, *Cunlin Wu, Yang Xiao, Boshen Zhang, Mingyang Zhang, Zhiguo Cao, Joey Tianyi Zhou*
 - DeciWatch: A Simple Baseline for 10x Efficient 2D and 3D Pose Estimation, *Ailing Zeng, Xuan Ju, Lei Yang, Ruiyuan Gao, Xizhou Zhu, Bo Dai, Qiang Xu*
 - SmoothNet: A Plug-and-Play Network for Refining Human Poses in Videos, *Ailing Zeng, Lei Yang, Xuan Ju, Jiefeng Li, Jianyi Wang, Qiang Xu*
 - PoseTrans: A Simple yet Effective Pose Transformation Augmentation for Human Pose Estimation, *Wentao Jiang, Sheng Jin, Wentao Liu, Chen Qian, Ping Luo, Si Liu*
 - Overlooked Poses Actually Make Sense: Distilling Privileged Knowledge for Human Motion Prediction, *Xiaoning Sun, Qiongjie Cui, Huaijiang Sun, Bin Li, Weiqing Li, Jianfeng Lu*
 - Structural Triangulation: A Closed-Form Solution to Constrained 3D Human Pose Estimation, *Zhuo Chen, Xu Zhao, Xiaoyue Wan*
 - Audio-Driven Stylized Gesture Generation with Flow-Based Model, *Sheng Ye, Yu-Hui Wen, Yanan Sun, Ying He, Ziyang Zhang, Yaoyuan Wang, Weihua He, Yong-Jin Liu*
 - Self-Constrained Inference Optimization on Structural Groups for Human Pose Estimation, *Zhehan Kan, Shuoshuo Chen, Zeng Li, Zhihai He*
 - Skeleton-Parted Graph Scattering Networks for 3D Human Motion Prediction, *Maosen Li, Siheng Chen, Zijiang Zhang, Lingxi Xie, Qi Tian, Ya Zhang*
 - Rethinking Keypoint Representations: Modeling Keypoints and Poses as Objects for Multi-Person Human Pose Estimation, *William McNally, Kanav Vats, Alexander Wong, John McPhee*
 - VirtualPose: Learning Generalizable 3D Human Pose Models from Virtual Data, *Jiajun Su, Chunyu Wang, Xiaoxuan Ma, Wenjun Zeng, Yizhou Wang*
 - Poseur: Direct Human Pose Regression with Transformers, *Weian Mao, Yongtao Ge, Chunhua Shen, Zhi Tian, Xinlong Wang, Zhibin Wang, Anton van den Hengel*

- Regularizing Vector Embedding in Bottom-Up Human Pose Estimation, *Haixin Wang, Lu Zhou, Yingying Chen, Ming Tang, Jinqiao Wang*
- A Visual Navigation Perspective for Category-Level Object Pose Estimation, *Jiaxin Guo, Fangxun Zhong, Rong Xiong, Yun-Hui Liu, Yue Wang, Yiyi Liao*
- Faster VoxelPose: Real-Time 3D Human Pose Estimation by Orthographic Projection, *Hang Ye, Wentao Zhu, Chunyu Wang, Rujie Wu, Yizhou Wang*
- Neural Capture of Animatable 3D Human from Monocular Video, *Gusi Te, Xiu Li, Xiao Li, Jinglu Wang, Wei Hu, Yan Lu*
- General Object Pose Transformation Network from Unpaired Data, *Yukun Su, Guosheng Lin, Ruizhou Sun, Qingyao Wu*
- 3D Interacting Hand Pose Estimation by Hand De-Occlusion and Removal, *Hao Meng, Sheng Jin, Wentao Liu, Chen Qian, Mengxiang Lin, Wanli Ouyang, Ping Luo*
- DH-AUG: DH Forward Kinematics Model Driven Augmentation for 3D Human Pose Estimation, *Linzhi Huang, Jiahao Liang, Weihong Deng*

Computational Photography

- Boosting Event Stream Super-Resolution with a Recurrent Neural Network, *Wenming Weng, Yueyi Zhang, Zhiwei Xiong*
- Projective Parallel Single-Pixel Imaging to Overcome Global Illumination in 3D Structure Light Scanning, *Yuxi Li, Huijie Zhao, Hongzhi Jiang, Xudong Li*
- Semantic-Sparse Colorization Network for Deep Exemplar-Based Colorization, *Yunpeng Bai, Chao Dong, Zenghao Chai, Andong Wang, Zhengzhuo Xu, Chun Yuan*
- FAST-VQA: Efficient End-to-End Video Quality Assessment with Fragment Sampling, *Haoning Wu, Chaofeng Chen, Jingwen Hou, Liang Liao, Annan Wang, Wenxiu Sun, Qiong Yan, Weisi Lin*
- LEDNet: Joint Low-Light Enhancement and Deblurring in the Dark, *Shangchen Zhou, Chongyi Li, Chen Change Loy*
- Real-rawVSR: Real-World Raw Video Super-Resolution with a Benchmark Dataset, *Huanjing Yue, Zhiming Zhang, Jingyu Yang*
- NEST: Neural Event Stack for Event-Based Image Enhancement, *Mingui Teng, Chu Zhou, Hanyue Lou, Boxin Shi*
- Seeing Far in the Dark with Patterned Flash, *Zhanghao Sun, Jian Wang, Yicheng Wu, Shree Nayar*
- Simple Baselines for Image Restoration, *Liangyu Chen, Xiaojie Chu, Xiangyu Zhang, Jian Sun*
- Spike Transformer: Monocular Depth Estimation for Spiking Camera, *Jiyuan Zhang, Lulu Tang, Zhaoifei Yu, Jiwen Lu, Tiejun Huang*
- Improving Image Restoration by Revisiting Global Information Aggregation, *Xiaojie Chu, Liangyu Chen, Chengpeng Chen, Xin Lu*
- Data Association between Event Streams and Intensity Frames under Diverse Baselines, *Dehao Zhang, Qiankun Ding, Peiqi Duan, Chu Zhou, Boxin Shi*
- D2HNet: Joint Denoising and Deblurring with Hierarchical Network for Robust Night Image Restoration, *Yuzhi Zhao, Yongzhe Xu, Qiong Yan, Dingdong Yang, Xuehui Wang, Lai-Man Po*
- Learning Graph Neural Networks for Image Style Transfer, *Yongcheng Jing, Yining Mao, Yiding Yang, Yibing Zhan, Mingli Song, Xinchao Wang, Dacheng Tao*
- DeepPS2: Revisiting Photometric Stereo Using Two Differently Illuminated Images, *Ashish Tiwari, Shanmuganathan Raman*
- Instance Contour Adjustment via Structure-Driven CNN, *Shuchen Weng, Yi Wei, Ming-Ching Chang, Boxin Shi*

- Human-Centric Image Cropping with Partition-Aware and Content-Preserving Features, *Bo Zhang, Li Niu, Xing Zhao, Liqing Zhang*
- Video Interpolation by Event-Driven Anisotropic Adjustment of Optical Flow, *Song Wu, Kaichao You, Weihua He, Chen Yang, Yang Tian, Yaoyuan Wang, Ziyang Zhang, Jianxing Liao*
- Spatial-Separated Curve Rendering Network for Efficient and High-Resolution Image Harmonization, *Jingtang Liang, Xiaodong Cun, Chi-Man Pun, Jue Wang*
- Deep Semantic Statistics Matching (D2SM) Denoising Network, *Kangfu Mei, Vishal M. Patel, Rui Huang*
- Neural Space-Filling Curves, *Hanyu Wang, Kamal Gupta, Larry Davis, Abhinav Shrivastava*
- Seeing through a Black Box: Toward High-Quality Terahertz Imaging via Subspace-and-Attention Guided Restoration, *Weng-Tai Su, Yi-Chun Hung, Po-Jen Yu, Shang-Hua Yang, Chia-Wen Lin*
- PANDORA: Polarization-Aided Neural Decomposition of Radiance, *Akshat Dave, Yongyi Zhao, Ashok Veeraraghavan*

Datasets and Evaluation

- Benchmarking Omni-Vision Representation through the Lens of Visual Realms, *Yuanhan Zhang, Zhenfei Yin, Jing Shao, Ziwei Liu*
- Neuromorphic Data Augmentation for Training Spiking Neural Networks, *Yuhang Li, Youngeun Kim, Hyoungeob Park, Tamar Geller, Priyadarshini Panda*
- CelebV-HQ: A Large-Scale Video Facial Attributes Dataset, *Hao Zhu, Wayne Wu, Wentao Zhu, Liming Jiang, Siwei Tang, Li Zhang, Ziwei Liu, Chen Change Loy*
- MovieCuts: A New Dataset and Benchmark for Cut Type Recognition, *Alejandro Pardo, Fabian Caba, Juan León Alcázar, Ali Thabet, Bernard Ghanem*
- Unitail: Detecting, Reading, and Matching in Retail Scene, *Fangyi Chen, Han Zhang, Zaiwang Li, Jiachen Dou, Shentong Mo, Hao Chen, Yongxin Zhang, Uzair Ahmed, Chenchen Zhu, Marios Savvides*
- REALY: Rethinking the Evaluation of 3D Face Reconstruction, *Zenghao Chai, Haoxian Zhang, Jing Ren, Di Kang, Zhengzhuo Xu, Xuefei Zhe, Chun Yuan, Linchao Bao*
- Capturing, Reconstructing, and Simulating: The UrbanScene3D Dataset, *Liqiang Lin, Yilin Liu, Yue Hu, Xingguang Yan, Ke Xie, Hui Huang*
- PartImageNet: A Large, High-Quality Dataset of Parts, *Ju He, Shuo Yang, Shaokang Yang, Adam Kortylewski, Xiaoding Yuan, Jie-Neng Chen, Shuai Liu, Cheng Yang, Qihang Yu, Alan Yuille*
- PANDORA: A Panoramic Detection Dataset for Object with Orientation, *Hang Xu, Qiang Zhao, Yike Ma, Xiaodong Li, Peng Yuan, Bailan Feng, Chenggang Yan, Feng Dai*
- MimicME: A Large Scale Diverse 4D Database for Facial Expression Analysis, *Athanasios Papaioannou, Baris Gecer, Shiyang Cheng, Grigorios G. Chrysos, Jiankang Deng, Eftychia Fotiadou, Christos Kampouris, Dimitrios Kollias, Stylianos Moschoglou, Kritaphat Songsri-In, Stylianos Ploumpis, George Trigeorgis, Panagiotis Tzirakis, Evangelos Ververas, Yuxiang Zhou, Allan Ponniah, Anastasios Roussos, Stefanos Zafeiriou*
- Delving into Universal Lesion Segmentation: Method, Dataset, and Benchmark, *Yu Qiu, Jing Xu*
- D2-TPred: Discontinuous Dependency for Trajectory Prediction under Traffic Lights, *Yuzhen Zhang, Wentong Wang, Weizhi Guo, Pei Lv, Mingliang Xu, Wei Chen, Dinesh Manocha*

- Learning Omnidirectional Flow in 360° Video via Siamese Representation, *Keshav Bhandari, Bin Duan, Gaowen Liu, Hugo Latapie, Ziliang Zong, Yan Yan*

Detection and Localization in 2D and/or 3D

- Deformable Feature Aggregation for Dynamic Multi-modal 3D Object Detection, *Zehui Chen, Zhenyu Li, Shiquan Zhang, Liangji Fang, Qinhong Jiang, Feng Zhao*
- MPPNet: Multi-Frame Feature Intertwining with Proxy Points for 3D Temporal Object Detection, *Xuesong Chen, Shaoshuai Shi, Benjin Zhu, Ka Chun Cheung, Hang Xu, Hongsheng Li*
- PTSEFormer: Progressive Temporal-Spatial Enhanced TransFormer towards Video Object Detection, *Han Wang, Jun Tang, Xiaodong Liu, Shanyan Guan, Rong Xie, Li Song*
- BEVFormer: Learning Bird's-Eye-View Representation from Multi-Camera Images via Spatiotemporal Transformers, *Zhiqi Li, Wenhai Wang, Hongyang Li, Enze Xie, Chonghao Sima, Tong Lu, Yu Qiao, Jifeng Dai*
- Category-Level 6D Object Pose and Size Estimation Using Self-Supervised Deep Prior Deformation Networks, *Jiehong Lin, Zewei Wei, Changxing Ding, Kui Jia*
- Dense Teacher: Dense Pseudo-Labels for Semi-Supervised Object Detection, *Hongyu Zhou, Zheng Ge, Songtao Liu, Weixin Mao, Zeming Li, Haiyan Yu, Jian Sun*
- Point-to-Box Network for Accurate Object Detection via Single Point Supervision, *Pengfei Chen, Xuehui Yu, Xumeng Han, Najmul Hassan, Kai Wang, Jiachen Li, Jian Zhao, Humphrey Shi, Zhenjun Han, Qixiang Ye*
- Towards Data-Efficient Detection Transformers, *Wen Wang, Jing Zhang, Yang Cao, Yongliang Shen, Dacheng Tao*
- Exploiting Unlabeled Data with Vision and Language Models for Object Detection, *Shiyu Zhao, Zhixing Zhang, Samuel Schulter, Long Zhao, Vijay Kumar B G, Anastasis Stathopoulos, Manmohan Chandraker, Dimitris N. Metaxas*
- INT: Towards Infinite-Frames 3D Detection with an Efficient Framework, *Jiayun Xu, Zhenwei Miao, Da Zhang, Hongyu Pan, Kaixuan Liu, Peihan Hao, Jun Zhu, Zhengyang Sun, Hongmin Li, Xin Zhan*
- End-to-End Weakly Supervised Object Detection with Sparse Proposal Evolution, *Mingxiang Liao, Fang Wan, Yuan Yao, Zhenjun Han, Jialing Zou, Yuze Wang, Bailan Feng, Peng Yuan, Qixiang Ye*
- Calibration-Free Multi-View Crowd Counting, *Qi Zhang, Antoni B. Chan*
- Unsupervised Domain Adaptation for Monocular 3D Object Detection via Self-Training, *Zhenyu Li, Zehui Chen, Ang Li, Liangji Fang, Qinhong Jiang, Xianming Liu, Junjun Jiang*
- SuperLine3D: Self-Supervised Line Segmentation and Description for LiDAR Point Cloud, *Xiangrui Zhao, Sheng Yang, Tianxin Huang, Jun Chen, Teng Ma, Mingyang Li, Yong Liu*
- Exploring Plain Vision Transformer Backbones for Object Detection, *Yanghao Li, Hanzi Mao, Ross Girshick, Kaiming He*
- HEAD: HEtero-Assists Distillation for Heterogeneous Object Detectors, *Luting Wang, Xiaojie Li, Yue Liao, Zeren Jiang, Jianlong Wu, Fei Wang, Chen Qian, Si Liu*
- You Should Look at All Objects, *Zhenchao Jin, Dongdong Yu, Luchuan Song, Zehuan Yuan, Lequan Yu*
- DCL-Net: Deep Correspondence Learning Network for 6D Pose Estimation, *Hongyang Li, Jiehong Lin, Kui Jia*
- DISP6D: Disentangled Implicit Shape and Pose Learning for Scalable 6D Pose Estimation, *Yilin Wen, Xiangyu Li, Hao Pan, Lei Yang, Zheng Wang, Taku Komura, Wenping Wang*
- Distilling Object Detectors with Global Knowledge, *Sanli Tang, Zhongyu Zhang, Zhanzhan Cheng, Jing Lu, Yunlu Xu, Yi Niu, Fan He*
- Unifying Visual Perception by Dispersible Points Learning, *Jianming Liang, Guanglu Song, Biao Leng, Yu Liu*
- PseCo: Pseudo Labeling and Consistency Training for Semi-Supervised Object Detection, *Gang Li, Xiang Li, Yujie Wang, Yichao Wu, Ding Liang, Shanshan Zhang*
- Robust Category-Level 6D Pose Estimation with Coarse-to-Fine Rendering of Neural Features, *Wufei Ma, Angtian Wang, Alan Yuille, Adam Kortylewski*
- Translation, Scale and Rotation: Cross-Modal Alignment Meets RGB-Infrared Vehicle Detection, *Maoxun Yuan, Yinyan Wang, Xingxing Wei*
- RFLA: Gaussian Receptive Field Based Label Assignment for Tiny Object Detection, *Chang Xu, Jinwang Wang, Wen Yang, Huai Yu, Lei Yu, Gui-Song Xia*
- TD-Road: Top-Down Road Network Extraction with Holistic Graph Construction, *Yang He, Ravi Garg, Amber Roy Chowdhury*
- Multi-faceted Distillation of Base-Novel Commonality for Few-Shot Object Detection, *Shuang Wu, Wenjie Pei, Dianwen Mei, Fanglin Chen, Jiandong Tian, Guangming Lu*
- PointCLM: A Contrastive Learning-Based Framework for Multi-Instance Point Cloud Registration, *Mingzhi Yuan, Zhihao Li, Qiuye Jin, Xinrong Chen, Manning Wang*
- Weakly Supervised Object Localization via Transformer with Implicit Spatial Calibration, *Haotian Bai, Ruimao Zhang, Jiong Wang, Xiang Wan*
- Label-Guided Auxiliary Training Improves 3D Object Detector, *Yaomin Huang, Xinmei Liu, Yichen Zhu, Zhiyuan Xu, Chaomin Shen, Zhengping Che, Guixu Zhang, Yaxin Peng, Feifei Feng, Jian Tang*
- PromptDet: Towards Open-Vocabulary Detection Using Uncurated Images, *Chengjian Feng, Yujie Zhong, Zequn Jie, Xiangxiang Chu, Haibing Ren, Xiaolin Wei, Weidi Xie, Lin Ma*
- Densely Constrained Depth Estimator for Monocular 3D Object Detection, *Yingyan Li, Yuntao Chen, Jiawei He, Zhaoxiang Zhang*
- Polarimetric Pose Prediction, *Daoyi Gao, Yitong Li, Patrick Ruhkamp, Iuliia Skobleva, Magdalena Wysocki, HyunJun Jung, Pengyuan Wang, Arturo Guridi, Benjamin Busam*
- Cornerformer: Purifying Instances for Corner-Based Detectors, *Haoran Wei, Xin Chen, Lingxi Xie, Qi Tian*
- PillarNet: Real-Time and High-Performance Pillar-Based 3D Object Detection, *Guangsheng Shi, Ruifeng Li, Chao Ma*
- Robust Object Detection with Inaccurate Bounding Boxes, *Chengxin Liu, Kewei Wang, Hao Lu, Zhiguo Cao, Ziming Zhang*
- Efficient Decoder-Free Object Detection with Transformers, *Peixian Chen, Mengdan Zhang, Yunhang Shen, Kekai Sheng, Yuting Gao, Xing Sun, Ke Li, Chunhua Shen*
- ReAct: Temporal Action Detection with Relational Queries, *Dingfeng Shi, Yujie Zhong, Qiong Cao, Jing Zhang, Lin Ma, Jia Li, Dacheng Tao*
- Towards Accurate Active Camera Localization, *Qihang Fang, Yingda Yin, Qingnan Fan, Fei Xia, Siyan Dong, Sheng Wang, Jue Wang, Leonidas J. Guibas, Baoquan Chen*
- Bagging Regional Classification Activation Maps for Weakly Supervised Object Localization, *Lei Zhu, Qian Chen, Lujia Jin, Yunfei You, Yanye Lu*
- UC-OWOD: Unknown-Classified Open World Object Detection, *Zhiheng Wu, Yue Lu, Xingyu Chen, Zhengxing Wu, Liwen Kang, Junzhi Yu*

- GTCaR: Graph Transformer for Camera Re-Localization, *Xinyi Li, Haibin Ling*
 - Few-Shot Object Detection by Knowledge Distillation Using Bag-of-Visual-Words Representations, *Wenjie Pei, Shuang Wu, Dianwen Mei, Fanglin Chen, Jiandong Tian, Guangming Lu*
 - SALISA: Saliency-Based Input Sampling for Efficient Video Object Detection, *Babak Ehteshami Bejnordi, Amirhossein Habibian, Fatih Porikli, Amir Ghodrati*
 - ECO-TR: Efficient Correspondences Finding via Coarse-to-Fine Refinement, *Dongli Tan, Jiang-Jiang Liu, Xingyu Chen, Chao Chen, Ruixin Zhang, Yunhang Shen, Shouhong Ding, Rongrong Ji*
 - DetMatch: Two Teachers Are Better than One for Joint 2D and 3D Semi-Supervised Object Detection, *Jinhyung Park, Chenfeng Xu, Yiyang Zhou, Masayoshi Tomizuka, Wei Zhan*
 - PCR-CG: Point Cloud Registration via Deep Explicit Color and Geometry, *Yu Zhang, Junle Yu, Xiaolin Huang, Wenhui Zhou, Ji Hou*
 - Video Anomaly Detection by Solving Decoupled Spatio-Temporal Jigsaw Puzzles, *Guodong Wang, Yunhong Wang, Jie Qin, Dongming Zhang, Xiuguo Bao, Di Huang*
 - Enhancing Multi-modal Features Using Local Self-Attention for 3D Object Detection, *Hao Li, Zehan Zhang, Xian Zhao, Yulong Wang, Yuxi Shen, Shiliang Pu, Hui Mao*
 - Weakly-Supervised Temporal Action Detection for Fine-Grained Videos with Hierarchical Atomic Actions, *Zhi Li, Lu He, Huijuan Xu*
 - Out-of-Distribution Identification: Let Detector Tell Which I Am Not Sure, *Ruoqi Li, Chongyang Zhang, Hao Zhou, Chao Shi, Yan Luo*
 - Learning with Free Object Segments for Long-Tailed Instance Segmentation, *Cheng Zhang, Tai-Yu Pan, Tianle Chen, Jike Zhong, Wenjin Fu, Wei-Lun Chao*
 - 3D Random Occlusion and Multi-layer Projection for Deep Multi-Camera Pedestrian Localization, *Rui Qiu, Ming Xu, Yuyao Yan, Jeremy S. Smith, Xi Yang*
 - A Simple Approach and Benchmark for 21,000-Category Object Detection, *Yutong Lin, Chen Li, Yue Cao, Zheng Zhang, Jianfeng Wang, Lijuan Wang, Zicheng Liu, Han Hu*
- Efficient Training and Inference Methods**
- Knowledge Condensation Distillation, *Chenxin Li, Mingbao Lin, Zhiyuan Ding, Nie Lin, Yihong Zhuang, Yue Huang, Xinghao Ding, Liujuan Cao*
 - Reducing Information Loss for Spiking Neural Networks, *Yufei Guo, Yuanpei Chen, Liwen Zhang, YingLei Wang, Xiaode Liu, Xinyi Tong, Yuanyuan Ou, Xuhui Huang, Zhe Ma*
 - Masked Generative Distillation, *Zhendong Yang, Zhe Li, Mingqi Shao, Dachuan Shi, Zehuan Yuan, Chun Yuan*
 - Fine-Grained Data Distribution Alignment for Post-Training Quantization, *Yunshan Zhong, Mingbao Lin, Mengzhao Chen, Ke Li, Yunhang Shen, Fei Chao, Yongjian Wu, Rongrong Ji*
 - Learning with Recoverable Forgetting, *Jingwen Ye, Yifang Fu, Jie Song, Xingyi Yang, Songhua Liu, Xin Jin, Mingli Song, Xinchao Wang*
 - Efficient One Pass Self-Distillation with Zipf's Label Smoothing, *Jiajun Liang, Linze Li, Zhaodong Bing, Borui Zhao, Yao Tang, Bo Lin, Haoqiang Fan*
 - Deep Partial Updating: Towards Communication Efficient Updating for On-Device Inference, *Zhongnan Qu, Cong Liu, Lothar Thiele*
 - Patch Similarity Aware Data-Free Quantization for Vision Transformers, *Zhikai Li, Liping Ma, Mengjuan Chen, Junrui Xiao, Qingyi Gu*
 - L3: Accelerator-Friendly Lossless Image Format for High-Resolution, High-Throughput DNN Training, *Jonghyun Bae, Woohyeon Baek, Tae Jun Ham, Jae W. Lee*
 - SP-Net: Slowly Progressing Dynamic Inference Networks, *Huanyu Wang, Wenhui Zhang, Shihao Su, Hui Wang, Zhenwei Miao, Xin Zhan, Xi Li*
 - Equivariance and Invariance Inductive Bias for Learning from Insufficient Data, *Tan Wang, Qianru Sun, Sugiri Pranata, Karlekar Jayashree, Hanwang Zhang*
 - Mixed-Precision Neural Network Quantization via Learned Layer-Wise Importance, *Chen Tang, Kai Ouyang, Zhi Wang, Yifei Zhu, Wen Ji, Yaowei Wang, Wenwu Zhu*
 - PalQuant: Accelerating High-Precision Networks on Low-Precision Accelerators, *Qinghao Hu, Gang Li, Qiman Wu, Jian Cheng*
 - Disentangled Differentiable Network Pruning, *Shangqian Gao, Feihu Huang, Yanfu Zhang, Heng Huang*
 - IDa-Det: An Information Discrepancy-Aware Distillation for 1-Bit Detectors, *Sheng Xu, Yanjing Li, Bohan Zeng, Teli Ma, Baochang Zhang, Xianbin Cao, Peng Gao, Jinhui Li*
 - Learning to Weight Samples for Dynamic Early-Exiting Networks, *Yizeng Han, Yifan Pu, Zihang Lai, Chaoqi Wang, Shiji Song, Junfeng Cao, Wenhui Huang, Chao Deng, Gao Huang*
 - AdaBin: Improving Binary Neural Networks with Adaptive Binary Sets, *Zhijun Tu, Xinghao Chen, Pengju Ren, Yunhe Wang*
 - Self-Slimmed Vision Transformer, *Zhuofan Zong, Kunchang Li, Guanglu Song, Yali Wang, Yu Qiao, Biao Leng, Yu Liu*
 - Switchable Online Knowledge Distillation, *Biao Qian, Yang Wang, Hongzhi Yin, Richang Hong, Meng Wang*
 - ∞ -Robustness and Beyond: Unleashing Efficient Adversarial Training, *Hadi M. Dolatabadi, Sarah Erfani, Christopher Leckie*
 - Multi-Granularity Pruning for Model Acceleration on Mobile Devices, *Tianli Zhao, Xi Sheryl Zhang, Wentao Zhu, Jiaxing Wang, Sen Yang, Ji Liu, Jian Cheng*
 - Towards Accurate Binary Neural Networks via Modeling Contextual Dependencies, *Xingrun Xing, Yangguang Li, Wei Li, Wenrui Ding, Yalong Jiang, Yufeng Wang, Jing Shao, Chunlei Liu, Xianglong Liu*
 - SPViT: Enabling Faster Vision Transformers via Latency-Aware Soft Token Pruning, *Zhenglun Kong, Peiyan Dong, Xiaolong Ma, Xin Meng, Wei Niu, Mengshu Sun, Xuan Shen, Geng Yuan, Bin Ren, Hao Tang, Minghai Qin, Yanzhi Wang*
 - Soft Masking for Cost-Constrained Channel Pruning, *Ryan Humble, Maying Shen, Jorge Albericio Latorre, Eric Darve, Jose Alvarez*
 - SuperTickets: Drawing Task-Agnostic Lottery Tickets from Supernet via Jointly Architecture Searching and Parameter Pruning, *Haoran You, Baopu Li, Zhanyi Sun, Xu Ouyang, Yingyan Lin*
 - Meta-GF: Training Dynamic-Depth Neural Networks Harmoniously, *Yi Sun, Jian Li, Xin Xu*
 - Explicit Model Size Control and Relaxation via Smooth Regularization for Mixed-Precision Quantization, *Vladimir Chikin, Kirill Solodskikh, Irina Zhelavskaya*
 - You Already Have It: A Generator-Free Low-Precision DNN Training Framework Using Stochastic Rounding, *Geng Yuan, Sung-En Chang, Qing Jin, Alec Lu, Yanyu Li, Yushu Wu, Zhenglun Kong, Yanyue Xie, Peiyan Dong, Minghai Qin, Xiaolong Ma, Xulong Tang, Zhenman Fang, Yanzhi Wang*

- Real Spike: Learning Real-Valued Spikes for Spiking Neural Networks, *Yufei Guo, Liwen Zhang, Yuanpei Chen, Xinyi Tong, Xiaode Liu, YingLei Wang, Xuhui Huang, Zhe Ma*
- LANA: Latency Aware Network Acceleration, *Pavlo Molchanov, Jimmy Hall, Hongxu Yin, Jan Kautz, Nicolo Fusi, Arash Vahdat*
- PTQ4ViT: Post-Training Quantization for Vision Transformers with Twin Uniform Quantization, *Zhihang Yuan, Chenhao Xue, Yiqi Chen, Qiang Wu, Guangyu Sun*

Explainable AI for CV

- Understanding the Dynamics of DNNs Using Graph Modularity, *Yao Lu, Wen Yang, Yunzhe Zhang, Zuohui Chen, Jinyin Chen, Qi Xuan, Zhen Wang, Xiaoni Yang*
- HIVE: Evaluating the Human Interpretability of Visual Explanations, *Sunnie S. Y. Kim, Nicole Meister, Vikram V. Ramaswamy, Ruth Fong, Olga Russakovsky*
- Contributions of Shape, Texture, and Color in Visual Recognition, *Yunhao Ge, Yao Xiao, Zhi Xu, Xingrui Wang, Laurent Itti*
- Are Vision Transformers Robust to Patch Perturbations? *Jindong Gu, Volker Tresp, Yao Qin*
- Shap-CAM: Visual Explanations for Convolutional Neural Networks Based on Shapley Value, *Quan Zheng, Ziwei Wang, Jie Zhou, Jiwen Lu*

Faces

- Privacy-Preserving Face Recognition with Learnable Privacy Budgets in Frequency Domain, *Jiazhen Ji, Huan Wang, Yuge Huang, Jiaxiang Wu, Xingkun Xu, Shouhong Ding, Shengchuan Zhang, Liujuan Cao, Rongrong Ji*
- Source-Free Domain Adaptation with Contrastive Domain Alignment and Self-Supervised Exploration for Face Anti-Spoofing, *Yuchen Liu, Yabo Chen, Wenrui Dai, Mengran Gou, Chun-Ting Huang, Hongkai Xiong*
- OneFace: One Threshold for All, *Jiaheng Liu, Zhipeng Yu, Haoyu Qin, Yichao Wu, Ding Liang, Gangming Zhao, Ke Xu*
- Label2Label: A Language Modeling Framework for Multi-Attribute Learning, *Wanhua Li, Zhexuan Cao, Jianjiang Feng, Jie Zhou, Jiwen Lu*
- AgeTransGAN for Facial Age Transformation with Rectified Performance Metrics, *Gee-Sern Hsu, Rui-Cang Xie, Zhi-Ting Chen, Yu-Hong Lin*
- Hierarchical Contrastive Inconsistency Learning for Deepfake Video Detection, *Zhihao Gu, Taiping Yao, Yang Chen, Shouhong Ding, Lizhuang Ma*
- Rethinking Robust Representation Learning under Fine-Grained Noisy Faces, *Bingqi Ma, Guanglu Song, Boxiao Liu, Yu Liu*
- Teaching Where to Look: Attention Similarity Knowledge Distillation for Low Resolution Face Recognition, *Sungho Shin, Joosoon Lee, Junseok Lee, Yeonguk Yu, Kyoobin Lee*
- Learning Dynamic Facial Radiance Fields for Few-Shot Talking Head Synthesis, *Shuai Shen, Wanhua Li, Zheng Zhu, Yueqi Duan, Jie Zhou, Jiwen Lu*
- CoupleFace: Relation Matters for Face Recognition Distillation, *Jiaheng Liu, Haoyu Qin, Yichao Wu, Jinyang Guo, Ding Liang, Ke Xu*
- Towards Unbiased Label Distribution Learning for Facial Pose Estimation Using Anisotropic Spherical Gaussian, *Zhiwen Cao, Dongfang Liu, Qifan Wang, Yingjie Chen*
- AU-Aware 3D Face Reconstruction through Personalized AU-Specific Blendshape Learning, *Chenyi Kuang, Zijun Cui, Jeffrey O. Kephart, Qiang Ji*

- BézierPalm: A Free Lunch for Palmprint Recognition, *Kai Zhao, Lei Shen, Yingyi Zhang, Chuhan Zhou, Tao Wang, Ruixin Zhang, Shouhong Ding, Wei Jia, Wei Shen*
- Adaptive Transformers for Robust Few-Shot Cross-Domain Face Anti-Spoofing, *Hsin-Ping Huang, Deqing Sun, Yaojie Liu, Wen-Sheng Chu, Taihong Xiao, Jinwei Yuan, Hartwig Adam, Ming-Hsuan Yang*
- Face2Face^P: Real-Time High-Resolution One-Shot Face Reenactment, *Kewei Yang, Kang Chen, Daoliang Guo, Song-Hai Zhang, Yuan-Chen Guo, Weidong Zhang*
- BoundaryFace: A Mining Framework with Noise Label Self-Correction for Face Recognition, *Shijie Wu, Xun Gong*

Fairness, Accountability, Transparency, and Ethics in Vision

- Discover and Mitigate Unknown Biases with Debiasing Alternate Networks, *Zhiheng Li, Anthony Hoogs, Chenliang Xu*
- Unsupervised and Semi-Supervised Bias Benchmarking in Face Recognition, *Alexandra Chouldechova, Siqi Deng, Yongxin Wang, Wei Xia, Pietro Perona*
- Towards Efficient Adversarial Training on Vision Transformers, *Boxi Wu, Jindong Gu, Zhifeng Li, Deng Cai, Xiaofei He, Wei Liu*
- MIME: Minority Inclusion for Majority Group Enhancement of AI Performance, *Pradyumna Chari, Yunhao Ba, Shreeram Athreya, Achuta Kadambi*
- An Invisible Black-Box Backdoor Attack through Frequency Domain, *Tong Wang, Yuan Yao, Feng Xu, Shengwei An, Hanghang Tong, Ting Wang*
- FairGRAPE: Fairness-Aware GRAdient Pruning mEthod for Face Attribute Classification, *Xiaofeng Lin, Seungbae Kim, Jungseock Joo*
- Attaining Class-Level Forgetting in Pretrained Model Using Few Samples, *Pravendra Singh, Pratik Mazumder, Mohammed Asad Karim*
- Anti-Neuron Watermarking: Protecting Personal Data against Unauthorized Neural Networks, *Zihang Zou, Boqing Gong, Liqiang Wang*
- Recover Fair Deep Classification Models via Altering Pre-trained Structure, *Yanfu Zhang, Shangqian Gao, Heng Huang*

First Person (Egocentric) Vision

- GIMO: Gaze-Informed Human Motion Prediction in Context, *Yang Zheng, Yanchao Yang, Kaichun Mo, Jiaman Li, Tao Yu, Yebin Liu, Karen Liu, Leonidas J. Guibas*

Image and Video Manipulation Detection

- Detecting and Recovering Sequential DeepFake Manipulation, *Rui Shao, Tianxing Wu, Ziwei Liu*
- Watermark Vaccine: Adversarial Attacks to Prevent Watermark Removal, *Xinwei Liu, Jian Liu, Yang Bai, Jindong Gu, Tao Chen, Xiaojun Jia, Xiaochun Cao*
- Explaining Deepfake Detection by Analysing Image Matching, *Shichao Dong, Jin Wang, Jiajun Liang, Haoqiang Fan, Renhe Ji*
- Detecting Generated Images by Real Images, *Bo Liu, Fan Yang, Xiuli Bi, Bin Xiao, Weisheng Li, Xinbo Gao*
- An Information Theoretic Approach for Attention-Driven Face Forgery Detection, *Ke Sun, Hong Liu, Taiping Yao, Xiaoshuai Sun, Shen Chen, Shouhong Ding, Rongrong Ji*
- Exploring Disentangled Content Information for Face Forgery Detection, *Jiahao Liang, Huafeng Shi, Weihong Deng*

Image and Video Retrieval

- Dual-Stream Knowledge-Preserving Hashing for Unsupervised Video Retrieval, *Pandeng Li, Hongtao Xie, Jiannan Ge, Lei Zhang, Shaobo Min, Yongdong Zhang*

- PASS: Part-Aware Self-Supervised Pre-training for Person Re-identification, *Kuan Zhu, Haiyun Guo, Tianyi Yan, Yousong Zhu, Jinqiao Wang, Ming Tang*
- Adaptive Cross-Domain Learning for Generalizable Person Re-identification, *Pengyi Zhang, Huanzhang Dou, Yunlong Yu, Xi Li*
- Multi-Query Video Retrieval, *Zeyu Wang, Yu Wu, Karthik Narasimhan, Olga Russakovsky*
- Learning Semantic Correspondence with Sparse Annotations, *Shuaiyi Huang, Luyu Yang, Bo He, Songyang Zhang, Xuming He, Abhinav Shrivastava*
- Dynamically Transformed Instance Normalization Network for Generalizable Person Re-identification, *Bingliang Jiao, Lingqiao Liu, Liying Gao, Guosheng Lin, Lu Yang, Shizhou Zhang, Peng Wang, Yanning Zhang*
- TS2-Net: Token Shift and Selection Transformer for Text-Video Retrieval, *Yuqi Liu, Pengfei Xiong, Luhui Xu, Shengming Cao, Qin Jin*
- Unstructured Feature Decoupling for Vehicle Re-identification, *Wen Qian, Hao Luo, Silong Peng, Fan Wang, Chen Chen, Hao Li*
- Mimic Embedding via Adaptive Aggregation: Learning Generalizable Person Re-identification, *Boqiang Xu, Jian Liang, Lingxiao He, Zhenan Sun*
- Granularity-Aware Adaptation for Image Retrieval over Multiple Tasks, *Jon Almazán, Byungsoo Ko, Geonmo Gu, Diane Larlus, Yannis Kalantidis*
- RVSL: Robust Vehicle Similarity Learning in Real Hazy Scenes Based on Semi-Supervised Learning, *Wei-Ting Chen, I-Hsiang Chen, Chih-Yuan Yeh, Hao-Hsiang Yang, Hua-En Chang, Jian-Jiun Ding, Sy-Yen Kuo*
- Lightweight Attentional Feature Fusion: A New Baseline for Text-to-Video Retrieval, *Fan Hu, Aozhu Chen, Ziyue Wang, Fangming Zhou, Jianfeng Dong, Xirong Li*
- Modality Synergy Complement Learning with Cascaded Aggregation for Visible-Infrared Person Re-identification, *Yiyuan Zhang, Sanyuan Zhao, Yuhao Kang, Jianbing Shen*
- Cross-Modality Transformer for Visible-Infrared Person Re-identification, *Kongzhu Jiang, Tianzhu Zhang, Xiang Liu, Bingqiao Qian, Yongdong Zhang, Feng Wu*
- Connecting Compression Spaces with Transformer for Approximate Nearest Neighbor Search, *Haokui Zhang, Buzhou Tang, Wenze Hu, Xiaoyu Wang*
- SEMICON: A Learning-to-Hash Solution for Large-Scale Fine-Grained Image Retrieval, *Yang Shen, Xuhao Sun, Xiu-Shen Wei, Qing-Yuan Jiang, Jian Yang*
- CAViT: Contextual Alignment Vision Transformer for Video Object Re-identification, *Jinlin Wu, Lingxiao He, Wu Liu, Yang Yang, Zhen Lei, Tao Mei, Stan Z. Li*
- Text-Based Temporal Localization of Novel Events, *Sudipta Paul, Niluthpol Chowdhury Mithun, Amit K. Roy-Chowdhury*
- Reliability-Aware Prediction via Uncertainty Learning for Person Image Retrieval, *Zhaopeng Dou, Zhongdao Wang, Weihua Chen, Yali Li, Shengjin Wang*
- **Image and Video Synthesis**
- Relighting4D: Neural Relightable Human from Videos, *Zhaoxi Chen, Ziwei Liu*
- Real-Time Intermediate Flow Estimation for Video Frame Interpolation, *Zhewei Huang, Tianyuan Zhang, Wen Heng, Boxin Shi, Shuchang Zhou*
- PixelFolder: An Efficient Progressive Pixel Synthesis Network for Image Generation, *Jing He, Yiyi Zhou, Qi Zhang, Jun Peng, Yunhang Shen, Xiaoshuai Sun, Chao Chen, Rongrong Ji*
- StyleSwap: Style-Based Generator Empowers Robust Face Swapping, *Zhiliang Xu, Hang Zhou, Zhibin Hong, Ziwei Liu, Jiaming Liu, Zhizhi Guo, Junyu Han, Jingtuo Liu, Errui Ding, Jingdong Wang*
- WaveGAN: Frequency-Aware GAN for High-Fidelity Few-Shot Image Generation, *Mengping Yang, Zhe Wang, Ziqiu Chi, Wenyi Feng*
- High-Fidelity GAN Inversion with Padding Space, *Qingyan Bai, Yinghao Xu, Jiapeng Zhu, Weihao Xia, Yujiu Yang, Yujun Shen*
- Designing One Unified Framework for High-Fidelity Face Reenactment and Swapping, *Chao Xu, Jiangning Zhang, Yue Han, Guanzhong Tian, Xianfang Zeng, Ying Tai, Yabiao Wang, Chengjie Wang, Yong Liu*
- Sobolev Training for Implicit Neural Representations with Approximated Image Derivatives, *Wentao Yuan, Qingtian Zhu, Xiangyue Liu, Yikang Ding, Haotian Zhang, Chi Zhang*
- Multi-Curve Translator for High-Resolution Photorealistic Image Translation, *Yuda Song, Hui Qian, Xin Du*
- Deep Bayesian Video Frame Interpolation, *Zhiyang Yu, Yu Zhang, Xujie Xiang, Dongqing Zou, Xijun Chen, Jimmy S. Ren*
- Cross Attention Based Style Distribution for Controllable Person Image Synthesis, *Xinyue Zhou, Mingyu Yin, Xinyuan Chen, Li Sun, Changxin Gao, Qingli Li*
- L-Tracing: Fast Light Visibility Estimation on Neural Surfaces by Sphere Tracing, *Ziyu Chen, Chenjing Ding, Jianfei Guo, Dongliang Wang, Yikang Li, Xuan Xiao, Wei Wu, Li Song*
- A Perceptual Quality Metric for Video Frame Interpolation, *Qiqi Hou, Abhijay Ghildyal, Feng Liu*
- Adaptive Feature Interpolation for Low-Shot Image Generation, *Mengyu Dai, Haibin Hang, Xiaoyang Guo*
- PalGAN: Image Colorization with Palette Generative Adversarial Networks, *Yi Wang, Menghan Xia, Lu Qi, Jing Shao, Yu Qiao*
- Fast-Vid2Vid: Spatial-Temporal Compression for Video-to-Video Synthesis, *Long Zhuo, Guangcong Wang, Shikai Li, Wayne Wu, Ziwei Liu*
- Learning Prior Feature and Attention Enhanced Image Inpainting, *Chenjie Cao, Qiaole Dong, Yanwei Fu*
- Temporal-MPI: Enabling Multi-Plane Images for Dynamic Scene Modelling via Temporal Basis Learning, *Wenpeng Xing, Jie Chen*
- 3D-Aware Semantic-Guided Generative Model for Human Synthesis, *Jichao Zhang, Enver Sangineto, Hao Tang, Aliaksandr Siarohin, Zhun Zhong, Nicu Sebe, Wei Wang*
- Scraping Textures from Natural Images for Synthesis and Editing, *Xueting Li, Xiaolong Wang, Ming-Hsuan Yang, Alexei A. Efros, Sifei Liu*
- Single Stage Virtual Try-On via Deformable Attention Flows, *Shuai Bai, Huiling Zhou, Zhikang Li, Chang Zhou, Hongxia Yang*
- StyleLight: HDR Panorama Generation for Lighting Estimation and Editing, *Guangcong Wang, Yinyu Yang, Chen Change Loy, Ziwei Liu*
- Learning Cross-Video Neural Representations for High-Quality Frame Interpolation, *Wentao Shangguan, Yu Sun, Weijie Gan, Ulugbek S. Kamilov*
- End-to-End Graph-Constrained Vectorized Floorplan Generation with Panoptic Refinement, *Jiachen Liu, Yuan Xue, Jose Duarte, Krishnendra Shekhawat, Zihan Zhou, Xiaolei Huang*
- FakeCLR: Exploring Contrastive Learning for Solving Latent Discontinuity in Data-Efficient GANs, *Ziqiang Li, Chaoyue Wang, Heliang Zheng, Jing Zhang, Bin Li*

- Unified Implicit Neural Stylization, *Zhiwen Fan, Yifan Jiang, Peihao Wang, Xinyu Gong, Dejie Xu, Zhangyang Wang*
 - GAN with Multivariate Disentangling for Controllable Hair Editing, *Xuyang Guo, Meina Kan, Tianle Chen, Shiguang Shan*
 - Harmonizer: Learning to Perform White-Box Image and Video Harmonization, *Zhanghan Ke, Chunyi Sun, Lei Zhu, Ke Xu, Rynson W.H. Lau*
 - Digging into Radiance Grid for Real-Time View Synthesis with Detail Preservation, *Jian Zhang, Jinchi Huang, Bowen Cai, Huan Fu, Mingming Gong, Chaohui Wang, Jiaming Wang, Hongchen Luo, Rongfei Jia, Binqiang Zhao, Xing Tang*
 - StyleGAN-Human: A Data-Centric Odyssey of Human Generation, *Jianglin Fu, Shikai Li, Yuming Jiang, Kwan-Yee Lin, Chen Qian, Chen Change Loy, Wayne Wu, Ziwei Liu*
 - ColorFormer: Image Colorization via Color Memory Assisted Hybrid-Attention Transformer, *Xiaozhong Ji, Boyuan Jiang, Donghao Luo, Guangpin Tao, Wenqing Chu, Zhifeng Xie, Chengjie Wang, Ying Tai*
 - EAGAN: Efficient Two-Stage Evolutionary Architecture Search for GANs, *Guohao Ying, Xin He, Bin Gao, Bo Han, Xiaowen Chu*
 - Multimodal Conditional Image Synthesis with Product-of-Experts GANs, *Xun Huang, Arun Mallya, Ting-Chun Wang, Ming-Yu Liu*
 - Bi-Level Feature Alignment for Versatile Image Translation and Manipulation, *Fangneng Zhan, Yingchen Yu, Rongliang Wu, Jiahui Zhang, Kaiwen Cui, Aoran Xiao, Shijian Lu, Chunyan Miao*
 - High-Fidelity Image Inpainting with GAN Inversion, *Yongsheng Yu, Libo Zhang, Heng Fan, Tiejian Luo*
 - DeltaGAN: Towards Diverse Few-Shot Image Generation with Sample-Specific Delta, *Yan Hong, Li Niu, Jianfu Zhang, Liqing Zhang*
 - StyleFace: Towards Identity-Disentangled Face Generation on Megapixels, *Yuchen Luo, Junwei Zhu, Keke He, Wenqing Chu, Ying Tai, Chengjie Wang, Junchi Yan*
 - BIPS: Bi-modal Indoor Panorama Synthesis via Residual Depth-Aided Adversarial Learning, *Changgyoon Oh, Wonjune Cho, Yujeong Chae, Daehee Park, Lin Wang, Kuk-Jin Yoon*
 - Geometry-Aware Single-Image Full-Body Human Relighting, *Chaonan Ji, Tao Yu, Kaiwen Guo, Jingxin Liu, Yebin Liu*
 - Multiview Regenerative Morphing with Dual Flows, *Chih-Jung Tsai, Cheng Sun, Hwann-Tzong Chen*
 - Motion and Appearance Adaptation for Cross-Domain Motion Transfer, *Borun Xu, Biao Wang, Jinhong Deng, Jiale Tao, Tiezheng Ge, Yuning Jiang, Wen Li, Lixin Duan*
 - Spatio-Temporal Deformable Attention Network for Video Deblurring, *Huicong Zhang, Haozhe Xie, Hongxun Yao*
 - HairNet: Hairstyle Transfer with Pose Changes, *Peihao Zhu, Rameen Abdal, John Femiani, Peter Wonka*
 - Unbiased Multi-Modality Guidance for Image Inpainting, *Yongsheng Yu, Dawei Du, Libo Zhang, Tiejian Luo*
 - Motion Transformer for Unsupervised Image Animation, *Jiale Tao, Biao Wang, Tiezheng Ge, Yuning Jiang, Wen Li, Lixin Duan*
 - EleGANt: Exquisite and Locally Editable GAN for Makeup Transfer, *Chenyu Yang, Wanrong He, Yingqing Xu, Yang Gao*
 - Editing Out-of-Domain GAN Inversion via Differential Activations, *Haorui Song, Yong Du, Tianyi Xiang, Junyu Dong, Jing Qin, Shengfeng He*
 - On the Robustness of Quality Measures for GANs, *Motasem Alfarra, Juan C. Pérez, Anna Frühstück, Philip H. S. Torr, Peter Wonka, Bernard Ghanem*
 - StyleHEAT: One-Shot High-Resolution Editable Talking Face Generation via Pre-trained StyleGAN, *Fei Yin, Yong Zhang, Xiaodong Cun, Mingdeng Cao, Yanbo Fan, Xuan Wang, Qingyan Bai, Baoyuan Wu, Jue Wang, Yujiu Yang*
 - Transformers As Meta-Learners for Implicit Neural Representations, *Yinbo Chen, Xiaolong Wang*
 - A Codec Information Assisted Framework for Efficient Compressed Video Super-Resolution, *Hengsheng Zhang, Xueyi Zou, Jiaming Guo, Youliang Yan, Rong Xie, Li Song*
 - Injecting 3D Perception of Controllable NeRF-GAN into StyleGAN for Editable Portrait Image Synthesis, *Jeong-gi Kwak, Yuanming Li, Dongsik Yoon, Donghyeon Kim, David Han, Hanseok Ko*
 - Selective TransHDR: Transformer-Based Selective HDR Imaging Using Ghost Region Mask, *Jou Won Song, Ye-In Park, Kyeongbo Kong, Jaeho Kwak, Suk-Ju Kang*
 - Learning Series-Parallel Lookup Tables for Efficient Image Super-Resolution, *Cheng Ma, Jingyi Zhang, Jie Zhou, Jiwen Lu*
 - GeoAug: Data Augmentation for Few-Shot NeRF with Geometry Constraints, *Di Chen, Yu Liu, Lianghua Huang, Bin Wang, Pan Pan*
 - Learning Object Placement via Dual-Path Graph Completion, *Siyuan Zhou, Liu Liu, Li Niu, Liqing Zhang*
 - BLT: Bidirectional Layout Transformer for Controllable Layout Generation, *Xiang Kong, Lu Jiang, Huiwen Chang, Han Zhang, Yuan Hao, Haifeng Gong, Irfan Essa*
 - Rayleigh EigenDirections (REDs): Nonlinear GAN Latent Space Traversals for Multidimensional Features, *Guha Balakrishnan, Raghudeep Gadde, Aleix Martinez, Pietro Perona*
 - Context-Consistent Semantic Image Editing with Style-Preserved Modulation, *Wuyang Luo, Su Yang, Hong Wang, Bo Long, Weishan Zhang*
 - Eliminating Gradient Conflict in Reference-Based Line-Art Colorization, *Zekun Li, Zhengyang Geng, Zhao Kang, Wenyu Chen, Yibo Yang*
 - Unsupervised Learning of Efficient Geometry-Aware Neural Articulated Representations, *Atsuhiko Noguchi, Xiao Sun, Stephen Lin, Tatsuya Harada*
- Low-Level and Physics-Based Vision**
- JPEG Artifacts Removal via Contrastive Representation Learning, *Xi Wang, Xueyang Fu, Yurui Zhu, Zheng-Jun Zha*
 - Unpaired Deep Image Dehazing Using Contrastive Disentanglement Learning, *Xiang Chen, Zhentao Fan, Pengpeng Li, Longgang Dai, Caihua Kong, Zhuoran Zheng, Yufeng Huang, Yufeng Li*
 - Efficient Long-Range Attention Network for Image Super-Resolution, *Xindong Zhang, Hui Zeng, Shi Guo, Lei Zhang*
 - FlowFormer: A Transformer Architecture for Optical Flow, *Zhaoyang Huang, Xiaoyu Shi, Chao Zhang, Qiang Wang, Ka Chun Cheung, Hongwei Qin, Jifeng Dai, Hongsheng Li*
 - Coarse-to-Fine Sparse Transformer for Hyperspectral Image Reconstruction, *Yuanhao Cai, Jing Lin, Xiaowan Hu, Haoqian Wang, Xin Yuan, Yulun Zhang, Radu Timofte, Luc Van Gool*
 - Learning Shadow Correspondence for Video Shadow Detection, *Xinpeng Ding, Jingwen Yang, Xiaowei Hu, Xiaomeng Li*
 - Metric Learning Based Interactive Modulation for Real-World Super-Resolution, *Chong Mou, Yanze Wu, Xintao Wang, Chao Dong, Jian Zhang, Ying Shan*

- Dynamic Dual Trainable Bounds for Ultra-Low Precision Super-Resolution Networks, *Yunshan Zhong, Mingbao Lin, Xunchao Li, Ke Li, Yunhang Shen, Fei Chao, Yongjian Wu, Rongrong Ji*
- OSFormer: One-Stage Camouflaged Instance Segmentation with Transformers, *Jialun Pei, Tianyang Cheng, Deng-Ping Fan, He Tang, Chuanbo Chen, Luc Van Gool*
- Highly Accurate Dichotomous Image Segmentation, *Xuebin Qin, Hang Dai, Xiaobin Hu, Deng-Ping Fan, Ling Shao, Luc Van Gool*
- Boosting Supervised Dehazing Methods via Bi-Level Patch Reweighting, *Xingyu Jiang, Hongkun Dou, Chengwei Fu, Bingquan Dai, Tianrun Xu, Yue Deng*
- Flow-Guided Transformer for Video Inpainting, *Kaidong Zhang, Jingjing Fu, Dong Liu*
- Shift-tolerant Perceptual Similarity Metric, *Abhijay Ghildyal, Feng Liu*
- Uncertainty Learning in Kernel Estimation for Multi-stage Blind Image Super-Resolution, *Zhenxuan Fang, Weisheng Dong, Xin Li, Jinjian Wu, Leida Li, Guangming Shi*
- Learning Spatio-Temporal Downsampling for Effective Video Upscaling, *Xiaoyu Xiang, Yapeng Tian, Vijay Rengarajan, Lucas D. Young, Bo Zhu, Rakesh Ranjan*
- SepLUT: Separable Image-Adaptive Lookup Tables for Real-Time Image Enhancement, *Canqian Yang, Meiguang Jin, Yi Xu, Rui Zhang, Ying Chen, Huaida Liu*
- Blind Image Decomposition, *Junlin Han, Weihao Li, Pengfei Fang, Chunyi Sun, Jie Hong, Mohammad Ali Armin, Lars Petersson, Hongdong Li*
- MuLUT: Cooperating Multiple Look-Up Tables for Efficient Image Super-Resolution, *Jiacheng Li, Chang Chen, Zhen Cheng, Zhiwei Xiong*
- Learning Spatiotemporal Frequency-Transformer for Compressed Video Super-Resolution, *Zhongwei Qiu, Huan Yang, Jianlong Fu, Dongmei Fu*
- Spatial-Frequency Domain Information Integration for Pan-Sharpening, *Man Zhou, Jie Huang, Keyu Yan, Hu Yu, Xueyang Fu, Aiping Liu, Xian Wei, Feng Zhao*
- Efficient Meta-Tuning for Content-Aware Neural Video Delivery, *Xiaoqi Li, Jiaming Liu, Shizun Wang, Cheng Lyu, Ming Lu, Yurong Chen, Anbang Yao, Yandong Guo, Shanghang Zhang*
- Reference-Based Image Super-Resolution with Deformable Attention Transformer, *Jiezhong Cao, Jingyun Liang, Kai Zhang, Yawei Li, Yulun Zhang, Wenguan Wang, Luc Van Gool*
- Local Color Distributions Prior for Image Enhancement, *Haoyuan Wang, Ke Xu, Rynson W.H. Lau*
- L-CoDer: Language-Based Colorization with Color-Object Decoupling Transformer, *Zheng Chang, Shuchen Weng, Yu Li, Si Li, Boxin Shi*
- From Face to Natural Image: Learning Real Degradation for Blind Image Super-Resolution, *Xiaoming Li, Chaofeng Chen, Xianhui Lin, Wangmeng Zuo, Lei Zhang*
- Towards Interpretable Video Super-Resolution via Alternating Optimization, *Jiezhong Cao, Jingyun Liang, Kai Zhang, Wenguan Wang, Qin Wang, Yulun Zhang, Hao Tang, Luc Van Gool*
- Fast and High Quality Image Denoising via Malleable Convolution, *Yifan Jiang, Bartłomiej Wronski, Ben Mildenhall, Jonathan T. Barron, Zhangyang Wang, Tianfan Xue*
- TAPE: Task-Agnostic Prior Embedding for Image Restoration, *Lin Liu, Lingxi Xie, Xiaopeng Zhang, Shanxin Yuan, Xiangyu Chen, Wengang Zhou, Houqiang Li, Qi Tian*
- Uncertainty Inspired Underwater Image Enhancement, *Zhenqi Fu, Wu Wang, Yue Huang, Xinghao Ding, Kai-Kuang Ma*
- Hourglass Attention Network for Image Inpainting, *Ye Deng, Siqi Hui, Rongye Meng, Sanping Zhou, Jinjun Wang*
- Unfolded Deep Kernel Estimation for Blind Image Super-Resolution, *Hongyi Zheng, Hongwei Yong, Lei Zhang*
- ReCoNet: Recurrent Correction Network for Fast and Efficient Multi-Modality Image Fusion, *Zhanbo Huang, Jinyuan Liu, Xin Fan, Risheng Liu, Wei Zhong, Zhongxuan Luo*
- Content Adaptive Latents and Decoder for Neural Image Compression, *Guanbo Pan, Guo Lu, Zhihao Hu, Dong Xu*
- Efficient and Degradation-Adaptive Network for Real-World Image Super-Resolution, *Jie Liang, Hui Zeng, Lei Zhang*
- Unidirectional Video Denoising by Mimicking Backward Recurrent Modules with Look-Ahead Forward Ones, *Junyi Li, Xiaohe Wu, Zhenxing Niu, Wangmeng Zuo*
- Self-Supervised Learning for Real-World Super-Resolution from Dual Zoomed Observations, *Zhilu Zhang, Ruohao Wang, Hongzhi Zhang, Yunjin Chen, Wangmeng Zuo*
- Towards Efficient and Scale-Robust Ultra-High-Definition Image Demoiréing, *Xin Yu, Peng Dai, Wenbo Li, Lan Ma, Jiajun Shen, Jia Li, Xiaojuan Qi*
- ERDN: Equivalent Receptive Field Deformable Network for Video Deblurring, *Bangrui Jiang, Zhihui Xie, Zhen Xia, Songnan Li, Shan Liu*
- ART-SS: An Adaptive Rejection Technique for Semi-Supervised Restoration for Adverse Weather-Affected Images, *Rajeev Yasarla, Carey E. Priebe, Vishal M. Patel*
- Fusion from Decomposition: A Self-Supervised Decomposition Approach for Image Fusion, *Pengwei Liang, Junjun Jiang, Xianming Liu, Jiayi Ma*
- Neural Color Operators for Sequential Image Retouching, *Yili Wang, Xin Li, Kun Xu, Dongliang He, Qi Zhang, Fu Li, Errui Ding*
- Compiler-Aware Neural Architecture Search for On-Mobile Real-Time Super-Resolution, *Yushu Wu, Yifan Gong, Pu Zhao, Yanyu Li, Zheng Zhan, Wei Niu, Hao Tang, Minghai Qin, Bin Ren, Yanzhi Wang*
- Frequency and Spatial Dual Guidance for Image Dehazing, *Hu Yu, Naishan Zheng, Man Zhou, Jie Huang, Zeyu Xiao, Feng Zhao*
- Towards Real-World HDRTV Reconstruction: A Data Synthesis-Based Approach, *Zhen Cheng, Tao Wang, Yong Li, Fenglong Song, Chang Chen, Zhiwei Xiong*
- ARM: Any-Time Super-Resolution Method, *Bohong Chen, Mingbao Lin, Kekai Sheng, Mengdan Zhang, Peixian Chen, Ke Li, Liujuan Cao, Rongrong Ji*
- Attention-Aware Learning for Hyperparameter Prediction in Image Processing Pipelines, *Haina Qin, Longfei Han, Juan Wang, Congxuan Zhang, Yanwei Li, Bing Li, Weiming Hu*
- Memory-Augmented Model-Driven Network for Pansharpening, *Keyu Yan, Man Zhou, Li Zhang, Chengjun Xie*
- All You Need Is RAW: Defending against Adversarial Attacks with Camera Image Pipelines, *Yuxuan Zhang, Bo Dong, Felix Heide*
- Ghost-Free High Dynamic Range Imaging with Context-Aware Transformer, *Zhen Liu, Yinglong Wang, Bing Zeng, Shuaicheng Liu*
- Style-Guided Shadow Removal, *Jin Wan, Hui Yin, Zhenyao Wu, Xinyi Wu, Yanting Liu, Song Wang*
- D2C-SR: A Divergence to Convergence Approach for Real-World Image Super-Resolution, *Youwei Li, Haibin Huang, Lanpeng Jia, Haoqiang Fan, Shuaicheng Liu*

- Efficient Video Deblurring Guided by Motion Magnitude, *Yusheng Wang, Yunfan Lu, Ye Gao, Lin Wang, Zhihang Zhong, Yinqiang Zheng, Atsushi Yamashita*
- Single Frame Atmospheric Turbulence Mitigation: A Benchmark Study and a New Physics-Inspired Transformer Model, *Zhiyuan Mao, Ajay Jaiswal, Zhangyang Wang, Stanley H. Chan*
- Image Super-Resolution with Deep Dictionary, *Shunta Maeda*
- TempFormer: Temporally Consistent Transformer for Video Denoising, *Mingyang Song, Yang Zhang, Tunç O. Aydin*
- DRCNet: Dynamic Image Restoration Contrastive Network, *Fei Li, Lingfeng Shen, Yang Mi, Zhenbo Li*
- Transformer with Implicit Edges for Particle-Based Physics Simulation, *Yidi Shao, Chen Change Loy, Bo Dai*
- Rethinking Video Rain Streak Removal: A New Synthesis Model and a Deraining Network with Video Rain Prior, *Shuai Wang, Lei Zhu, Huazhu Fu, Jing Qin, Carola-Bibiane Schönlieb, Wei Feng, Song Wang*
- Super-Resolution by Predicting Offsets: An Ultra-Efficient Super-Resolution Network for Rasterized Images, *Jinjin Gu, Haoming Cai, Chenyu Dong, Ruofan Zhang, Yulun Zhang, Wenming Yang, Chun Yuan*
- Animation from Blur: Multi-modal Blur Decomposition with Motion Guidance, *Zhihang Zhong, Xiao Sun, Zhirong Wu, Yinqiang Zheng, Stephen Lin, Imari Sato*
- AlphaVC: High-Performance and Efficient Learned Video Compression, *Yibo Shi, Yunying Ge, Jing Wang, Jue Mao*
- Content-Oriented Learned Image Compression, *Meng Li, Shangyin Gao, Yihui Feng, Yibo Shi, Jing Wang*

Low-Shot Learning

- Contrastive Prototypical Network with Wasserstein Confidence Penalty, *Haoqing Wang, Zhi-Hong Deng*
- Learn-to-Decompose: Cascaded Decomposition Network for Cross-Domain Few-Shot Facial Expression Recognition, *Xinyi Zou, Yan Yan, Jing-Hao Xue, Si Chen, Hanzi Wang*
- Self-Support Few-Shot Semantic Segmentation, *Qi Fan, Wenjie Pei, Yu-Wing Tai, Chi-Keung Tang*
- Few-Shot Object Detection with Model Calibration, *Qi Fan, Chi-Keung Tang, Yu-Wing Tai*
- Self-Supervision Can Be a Good Few-Shot Learner, *Yuning Lu, Liangjian Wen, Jianzhuang Liu, Yajing Liu, Xinmei Tian*
- tSF: Transformer-Based Semantic Filter for Few-Shot Learning, *Jinxiang Lai, Siqian Yang, Wenlong Liu, Yi Zeng, Zhongyi Huang, Wenlong Wu, Jun Liu, Bin-Bin Gao, Chengjie Wang*
- Adversarial Feature Augmentation for Cross-Domain Few-Shot Classification, *Yanxu Hu, Andy J. Ma*
- Constructing Balance from Imbalance for Long-Tailed Image Recognition, *Yue Xu, Yong-Lu Li, Jiefeng Li, Cewu Lu*
- On Multi-Domain Long-Tailed Recognition, Imbalanced Domain Generalization and Beyond, *Yuzhe Yang, Hao Wang, Dina Katabi*
- Few-Shot Video Object Detection, *Qi Fan, Chi-Keung Tang, Yu-Wing Tai*
- Worst Case Matters for Few-Shot Recognition, *Minghao Fu, Yun-Hao Cao, Jianxin Wu*
- Exploring Hierarchical Graph Representation for Large-Scale Zero-Shot Image Classification, *Kai Yi, Xiaoqian Shen, Yunhao Gou, Mohamed Elhoseiny*
- Dense Cross-Query-and-Support Attention Weighted Mask Aggregation for Few-Shot Segmentation, *Xinyu Shi, Dong Wei,*

- Yu Zhang, Donghuan Lu, Munan Ning, Jiashun Chen, Kai Ma, Yefeng Zheng*
- Rethinking Clustering-Based Pseudo-Labeling for Unsupervised Meta-Learning, *Xingping Dong, Jianbing Shen, Ling Shao*
- Few-Shot Class-Incremental Learning for 3D Point Cloud Objects, *Townim Chowdhury, Ali Cheraghian, Sameera Ramasinghe, Sahar Ahmadi, Morteza Saberi, Shafin Rahman*
- Meta-Learning with Less Forgetting on Large-Scale Non-stationary Task Distributions, *Zhenyi Wang, Li Shen, Le Fang, Qiuling Suo, Donglin Zhan, Tiehang Duan, Mingchen Gao*
- DNA: Improving Few-Shot Transfer Learning with Low-Rank Decomposition and Alignment, *Ziyu Jiang, Tianlong Chen, Xuxi Chen, Yu Cheng, Luowei Zhou, Lu Yuan, Ahmed Awadallah, Zhangyang Wang*
- Open-World Semantic Segmentation via Contrasting and Clustering Vision-Language Embedding, *Quande Liu, Youpeng Wen, Jianhua Han, Chunjing Xu, Hang Xu, Xiaodan Liang*
- Few-Shot Classification with Contrastive Learning, *Zhanyuan Yang, Jinghua Wang, Yingying Zhu*
- Time-rEversed diffusiON tEensor Transformer: A New TENET of Few-Shot Object Detection, *Shan Zhang, Naila Murray, Lei Wang, Piotr Koniusz*
- Self-Promoted Supervision for Few-Shot Transformer, *Bowen Dong, Pan Zhou, Shuicheng Yan, Wangmeng Zuo*
- Few-Shot Object Counting and Detection, *Thanh Nguyen, Chau Pham, Khoi Nguyen, Minh Hoai*
- Cross-Domain Cross-Set Few-Shot Learning via Learning Compact and Aligned Representations, *Wentao Chen, Zhang Zhang, Wei Wang, Liang Wang, Zilei Wang, Tieniu Tan*
- Mutually Reinforcing Structure with Proposal Contrastive Consistency for Few-Shot Object Detection, *Tianxue Ma, Mingwei Bi, Jian Zhang, Wang Yuan, Zhizhong Zhang, Yuan Xie, Shouhong Ding, Lizhuang Ma*
- Dual Contrastive Learning with Anatomical Auxiliary Supervision for Few-Shot Medical Image Segmentation, *Huisi Wu, Fangyan Xiao, Chongxin Liang*
- Tree Structure-Aware Few-Shot Image Classification via Hierarchical Aggregation, *Min Zhang, Siteng Huang, Wenbin Li, Donglin Wang*
- Inductive and Transductive Few-Shot Video Classification via Appearance and Temporal Alignments, *Khoi D. Nguyen, Quoc-Huy Tran, Khoi Nguyen, Binh-Son Hua, Rang Nguyen*
- TransVLAD: Focusing on Locally Aggregated Descriptors for Few-Shot Learning, *Haoquan Li, Laoming Zhang, Daoan Zhang, Lang Fu, Peng Yang, Jianguo Zhang*
- Kernel Relative-Prototype Spectral Filtering for Few-Shot Learning, *Tao Zhang, Wu Huang*

Machine Learning Architectures and Formulations

- CLOSE: Curriculum Learning on the Sharing Extent towards Better One-Shot NAS, *Zixuan Zhou, Xuefei Ning, Yi Cai, Jiashu Han, Yiping Deng, Yuhao Dong, Huazhong Yang, Yu Wang*
- Online Continual Learning with Contrastive Vision Transformer, *Zhen Wang, Liu Liu, Yajing Kong, Jiaxian Guo, Dacheng Tao*
- EAUTOdet: Efficient Architecture Search for Object Detection, *Xiaoxing Wang, Jiale Lin, Juanping Zhao, Xiaokang Yang, Junchi Yan*
- A Max-Flow Based Approach for Neural Architecture Search, *Chao Xue, Xiaoxing Wang, Junchi Yan, Chun-Guang Li*
- Convolutional Embedding Makes Hierarchical Vision Transformer Stronger, *Cong Wang, Hongmin Xu, Xiong Zhang, Li Wang, Zhitong Zheng, Haifeng Liu*

- Active Label Correction Using Robust Parameter Update and Entropy Propagation, *Kwang In Kim*
 - AMixer: Adaptive Weight Mixing for Self-Attention Free Vision Transformers, *Yongming Rao, Wenliang Zhao, Jie Zhou, Jiwen Lu*
 - TinyViT: Fast Pretraining Distillation for Small Vision Transformers, *Kan Wu, Jinnian Zhang, Houwen Peng, Mengchen Liu, Bin Xiao, Jianlong Fu, Lu Yuan*
 - ScaleNet: Searching for the Model to Scale, *Jiyang Xie, Xiu Su, Shan You, Zhanyu Ma, Fei Wang, Chen Qian*
 - Complementing Brightness Constancy with Deep Networks for Optical Flow Prediction, *Vincent Le Guen, Clément Rambour, Nicolas Thome*
 - ViTAS: Vision Transformer Architecture Search, *Xiu Su, Shan You, Jiyang Xie, Mingkai Zheng, Fei Wang, Chen Qian, Changshui Zhang, Xiaogang Wang, Chang Xu*
 - Black-Box Few-Shot Knowledge Distillation, *Dang Nguyen, Sunil Gupta, Kien Do, Svetha Venkatesh*
 - Filter Pruning via Feature Discrimination in Deep Neural Networks, *Zhiqiang He, Yaguan Qian, Yuqi Wang, Bin Wang, Xiaohui Guan, Zhaoquan Gu, Xiang Ling, Shaoning Zeng, Haijiang Wang, Wujie Zhou*
 - LA3: Efficient Label-Aware AutoAugment, *Mingjun Zhao, Shan Lu, Zixuan Wang, Xiaoli Wang, Di Niu*
 - Interpretations Steered Network Pruning via Amortized Inferred Saliency Maps, *Alireza Ganjdanesh, Shangqian Gao, Heng Huang*
 - BA-Net: Bridge Attention for Deep Convolutional Neural Networks, *Yue Zhao, Junzhou Chen, Zirui Zhang, Ronghui Zhang*
 - SAU: Smooth Activation Function Using Convolution with Approximate Identities, *Koushik Biswas, Sandeep Kumar, Shilpak Banerjee, Ashish Kumar Pandey*
- Medical, Biological, and Cell Microscopy**
- Check and Link: Pairwise Lesion Correspondence Guides Mammogram Mass Detection, *Ziwei Zhao, Dong Wang, Yihong Chen, Ziteng Wang, Liwei Wang*
 - Generalizable Medical Image Segmentation via Random Amplitude Mixup and Domain-Specific Image Restoration, *Ziqi Zhou, Lei Qi, Yinghuan Shi*
 - Auto-FedRL: Federated Hyperparameter Optimization for Multi-Institutional Medical Image Segmentation, *Pengfei Guo, Dong Yang, Ali Hatamizadeh, An Xu, Ziyue Xu, Wenqi Li, Can Zhao, Daguang Xu, Stephanie Harmon, Evrim Turkbey, Baris Turkbey, Bradford Wood, Francesca Patella, Elvira Stellato, Gianpaolo Carrafiello, Vishal M. Patel, Holger R. Roth*
 - Personalizing Federated Medical Image Segmentation via Local Calibration, *Jiacheng Wang, Yueming Jin, Liansheng Wang*
 - One-Shot Medical Landmark Localization by Edge-Guided Transform and Noisy Landmark Refinement, *Zihao Yin, Ping Gong, Chunyu Wang, Yizhou Yu, Yizhou Wang*
 - Med-DANet: Dynamic Architecture Network for Efficient Medical Volumetric Segmentation, *Wenxuan Wang, Chen Chen, Jing Wang, Sen Zha, Yan Zhang, Jiangyun Li*
 - ConCL: Concept Contrastive Learning for Dense Prediction Pre-training in Pathology Images, *Jiawei Yang, Hanbo Chen, Yuan Liang, Junzhou Huang, Lei He, Jianhua Yao*
 - CryoAI: Amortized Inference of Poses for Ab Initio Reconstruction of 3D Molecular Volumes from Real Cryo-EM Images, *Axel Levy, Frédéric Poitevin, Julien Martel, Youssef Nashed, Ariana Peck, Nina Miolane, Daniel Ratner, Mike Dunne, Gordon Wetzstein*
 - UniMiSS: Universal Medical Self-Supervised Learning via Breaking Dimensionality Barrier, *Yutong Xie, Jianpeng Zhang, Yong Xia, Qi Wu*
 - Semi-Supervised Keypoint Detector and Descriptor for Retinal Image Matching, *Jiazhen Liu, Xirong Li, Qijie Wei, Jie Xu, Dayong Ding*
 - Accurate Detection of Proteins in Cryo-Electron Tomograms from Sparse Labels, *Qinwen Huang, Ye Zhou, Hsuan-Fu Liu, Alberto Bartesaghi*
 - RadioTransformer: A Cascaded Global-Focal Transformer for Visual Attention-Guided Disease Classification, *Moinak Bhattacharya, Shubham Jain, Prateek Prasanna*
- Motion and Tracking**
- Learning Uncoupled-Modulation CVAE for 3D Action-Conditioned Human Motion Synthesis, *Chongyang Zhong, Lei Hu, Zihao Zhang, Shihong Xia*
 - ByteTrack: Multi-Object Tracking by Associating Every Detection Box, *Yifu Zhang, Peize Sun, Yi Jiang, Dongdong Yu, Fucheng Weng, Zehuan Yuan, Ping Luo, Wenyu Liu, Xinggang Wang*
 - Robust Multi-Object Tracking by Marginal Inference, *Yifu Zhang, Chunyu Wang, Xinggang Wang, Wenjun Zeng, Wenyu Liu*
 - CMT: Context-Matching-Guided Transformer for 3D Tracking in Point Clouds, *Zhiyang Guo, Yunyao Mao, Wengang Zhou, Min Wang, Houqiang Li*
 - Hierarchical Latent Structure for Multi-modal Vehicle Trajectory Forecasting, *Dooseop Choi, KyoungWook Min*
 - AiATrack: Attention in Attention for Transformer Visual Tracking, *Shenyuan Gao, Chunluan Zhou, Chao Ma, Xinggang Wang, Junsong Yuan*
 - Social ODE: Multi-agent Trajectory Forecasting with Neural Ordinary Differential Equations, *Song Wen, Hao Wang, Dimitris N. Metaxas*
 - Sequential Multi-View Fusion Network for Fast LiDAR Point Motion Estimation, *Gang Zhang, Xiaoyan Li, Zhenhua Wang*
 - E-Graph: Minimal Solution for Rigid Rotation with Extensibility Graphs, *Yanyan Li, Federico Tombari*
 - Point Cloud Compression with Range Image-Based Entropy Model for Autonomous Driving, *Sukai Wang, Ming Liu*
 - Joint Feature Learning and Relation Modeling for Tracking: A One-Stream Framework, *Botao Ye, Hong Chang, Bingpeng Ma, Shiguang Shan, Xilin Chen*
 - Aware of the History: Trajectory Forecasting with the Local Behavior Data, *Yiqi Zhong, Zhenyang Ni, Siheng Chen, Ulrich Neumann*
 - Optical Flow Training under Limited Label Budget via Active Learning, *Shuai Yuan, Xian Sun, Hannah Kim, Shuzhi Yu, Carlo Tomasi*
 - Hierarchical Feature Embedding for Visual Tracking, *Zhixiong Pi, Weitao Wan, Chong Sun, Changxin Gao, Nong Sang, Chen Li*
 - Social-Implicit: Rethinking Trajectory Prediction Evaluation and the Effectiveness of Implicit Maximum Likelihood Estimation, *Abduallah Mohamed, Deyao Zhu, Warren Vu, Mohamed Elhoseiny, Christian Claudel*
 - Large-Displacement 3D Object Tracking with Hybrid Non-local Optimization, *Xuhui Tian, Xinran Lin, Fan Zhong, Xueying Qin*
 - View Vertically: A Hierarchical Network for Trajectory Prediction via Fourier Spectrums, *Conghao Wong, Beihao Xia, Ziming Hong, Qinmu Peng, Wei Yuan, Qiong Cao, Yibo Yang, Xinge You*

- HVC-Net: Unifying Homography, Visibility, and Confidence Learning for Planar Object Tracking, *Haoxian Zhang, Yonggen Ling*

Neural Generative Models

- RamGAN: Region Attentive Morphing GAN for Region-Level Makeup Transfer, *Jianfeng Xiang, Junliang Chen, Wenshuang Liu, Xianxu Hou, Linlin Shen*
- SinNeRF: Training Neural Radiance Fields on Complex Scenes from a Single Image, *Dejia Xu, Yifan Jiang, Peihao Wang, Zhiwen Fan, Humphrey Shi, Zhangyang Wang*
- Entropy-Driven Sampling and Training Scheme for Conditional Diffusion Generation, *Guangcong Zheng, Shengming Li, Hui Wang, Taiping Yao, Yang Chen, Shouhong Ding, Xi Li*
- Accelerating Score-Based Generative Models with Preconditioned Diffusion Sampling, *Hengyuan Ma, Li Zhang, Xiatian Zhu, Jianfeng Feng*
- Learning to Generate Realistic LiDAR Point Clouds, *Vlas Zyrjanov, Xiyue Zhu, Shenlong Wang*
- Diverse Image Inpainting with Normalizing Flow, *Cairong Wang, Yiming Zhu, Chun Yuan*
- Outpainting by Queries, *Kai Yao, Penglei Gao, Xi Yang, Jie Sun, Rui Zhang, Kaizhu Huang*
- Geometry-Guided Progressive NeRF for Generalizable and Efficient Neural Human Rendering, *Mingfei Chen, Jianfeng Zhang, Xiangyu Xu, Lijuan Liu, Yujun Cai, Jiashi Feng, Shuicheng Yan*
- Subspace Diffusion Generative Models, *Bowen Jing, Gabriele Corso, Renato Berlinghieri, Tommi Jaakkola*
- DuelGAN: A Duel between Two Discriminators Stabilizes the GAN Training, *Jiaheng Wei, Minghao Liu, Jiahao Luo, Andrew Zhu, James Davis, Yang Liu*

Optimization and Learning Methods

- An Embedded Feature Whitening Approach to Deep Neural Network Optimization, *Hongwei Yong, Lei Zhang*
- Self-Supervised Learning of Visual Graph Matching, *Chang Liu, Shaofeng Zhang, Xiaokang Yang, Junchi Yan*
- Scalable Learning to Optimize: A Learned Optimizer Can Train Big Models, *Xuxi Chen, Tianlong Chen, Yu Cheng, Weizhu Chen, Ahmed Awadallah, Zhangyang Wang*
- R-DFCIL: Relation-Guided Representation Learning for Data-Free Class Incremental Learning, *Qiankun Gao, Chen Zhao, Bernard Ghanem, Jian Zhang*
- Predicting Is Not Understanding: Recognizing and Addressing Underspecification in Machine Learning, *Damien Teney, Maxime Peyrard, Ehsan Abbasnejad*
- Neural-Sim: Learning to Generate Training Data with NeRF, *Yunhao Ge, Harkirat Behl, Jiashu Xu, Suriya Gunasekar, Neel Joshi, Yale Song, Xin Wang, Laurent Itti, Vibhav Vineet*
- Bayesian Optimization with Clustering and Rollback for CNN Auto Pruning, *Hanwei Fan, Jiandong Mu, Wei Zhang*
- Continual Variational Autoencoder Learning via Online Cooperative Memorization, *Fei Ye, Adrian G. Bors*
- Learning to Learn with Smooth Regularization, *Yuanhao Xiong, Cho-Jui Hsieh*
- Incremental Task Learning with Incremental Rank Updates, *Rakib Hyder, Ken Shao, Boyu Hou, Panos Markopoulos, Ashley Prater-Bennette, M. Salman Asif*
- Batch-Efficient EigenDecomposition for Small and Medium Matrices, *Yue Song, Nicu Sebe, Wei Wang*

- Ensemble Learning Priors Driven Deep Unfolding for Scalable Video Snapshot Compressive Imaging, *Chengshuai Yang, Shiyu Zhang, Xin Yuan*
- Approximate Discrete Optimal Transport Plan with Auxiliary Measure Method, *Dongsheng An, Na Lei, Xianfeng Gu*
- AdaBest: Minimizing Client Drift in Federated Learning via Adaptive Bias Estimation, *Farshid Varno, Marzie Saghay, Laya Rafiee Sevyeri, Sharut Gupta, Stan Matwin, Mohammad Havaei*

Recognition and Classification

- Doubly-Fused ViT: Fuse Information from Vision Transformer Doubly with Local Representation, *Li Gao, Dong Nie, Bo Li, Xiaofeng Ren*
- Improving Vision Transformers by Revisiting High-Frequency Components, *Jiawang Bai, Li Yuan, Shu-Tao Xia, Shuicheng Yan, Zhifeng Li, Wei Liu*
- Where to Focus: Investigating Hierarchical Attention Relationship for Fine-Grained Visual Classification, *Yang Liu, Lei Zhou, Pengcheng Zhang, Xiao Bai, Lin Gu, Xiaohan Yu, Jun Zhou, Edwin R. Hancock*
- DaViT: Dual Attention Vision Transformers, *Mingyu Ding, Bin Xiao, Noel Codella, Ping Luo, Jingdong Wang, Lu Yuan*
- Optimal Transport for Label-Efficient Visible-Infrared Person Re-identification, *Jiangming Wang, Zhizhong Zhang, Mingang Chen, Yi Zhang, Cong Wang, Bin Sheng, Yanyun Qu, Yuan Xie*
- Locality Guidance for Improving Vision Transformers on Tiny Datasets, *Kehan Li, Runyi Yu, Zhennan Wang, Li Yuan, Guoli Song, Jie Chen*
- Neighborhood Collective Estimation for Noisy Label Identification and Correction, *Jichang Li, Guanbin Li, Feng Liu, Yizhou Yu*
- Few-Shot Class-Incremental Learning via Entropy-Regularized Data-Free Replay, *Huan Liu, Li Gu, Zhixiang Chi, Yang Wang, Yuanhao Yu, Jun Chen, Jin Tang*
- Anti-Retroactive Interference for Lifelong Learning, *Runqi Wang, Yuxiang Bao, Baochang Zhang, Jianzhuang Liu, Wentao Zhu, Guodong Guo*
- Dynamic Metric Learning with Cross-Level Concept Distillation, *Wenzhao Zheng, Yuanhui Huang, Borui Zhang, Jie Zhou, Jiwen Lu*
- MENet: A Memory-Based Network with Dual-Branch for Efficient Event Stream Processing, *Linhui Sun, Yifan Zhang, Ke Cheng, Jian Cheng, Hanqing Lu*
- Out-of-Distribution Detection with Boundary Aware Learning, *Sen Pei, Xin Zhang, Bin Fan, Gaofeng Meng*
- KVT: k-NN Attention for Boosting Vision Transformers, *Pichao Wang, Xue Wang, Fan Wang, Ming Lin, Shuning Chang, Hao Li, Rong Jin*
- Learning Invariant Visual Representations for Compositional Zero-Shot Learning, *Tian Zhang, Kongming Liang, Ruoyi Du, Xian Sun, Zhanyu Ma, Jun Guo*
- Improving Covariance Conditioning of the SVD Meta-Layer by Orthogonality, *Yue Song, Nicu Sebe, Wei Wang*
- Out-of-Distribution Detection with Semantic Mismatch under Masking, *Yijun Yang, Ruiyuan Gao, Qiang Xu*
- Learning from Multiple Annotator Noisy Labels via Sample-Wise Label Fusion, *Zhengqi Gao, Fan-Keng Sun, Mingran Yang, Sucheng Ren, Zikai Xiong, Marc Engeler, Antonio Burazer, Linda Wildling, Luca Daniel, Duane S. Boning*
- Acknowledging the Unknown for Multi-Label Learning with Single Positive Labels, *Donghao Zhou, Pengfei Chen, Qiong Wang, Guangyong Chen, Pheng-Ann Heng*

- MaxViT: Multi-axis Vision Transformer, *Zhengzhong Tu, Hossein Talebi, Han Zhang, Feng Yang, Peyman Milanfar, Alan Bovik, Yinxiao Li*
 - ScalableViT: Rethinking the Context-Oriented Generalization of Vision Transformer, *Rui Yang, Hailong Ma, Jie Wu, Yansong Tang, Xuefeng Xiao, Min Zheng, Xiu Li*
 - MixSKD: Self-Knowledge Distillation from Mixup for Image Recognition, *Chuangang Yang, Zhulin An, Helong Zhou, Linhang Cai, Xiang Zhi, Jiwen Wu, Yongjun Xu, Qian Zhang*
 - Self-Feature Distillation with Uncertainty Modeling for Degraded Image Recognition, *Zhou Yang, Weisheng Dong, Xin Li, Jinjian Wu, Leida Li, Guangming Shi*
 - Novel Class Discovery without Forgetting, *K J Joseph, Sujoy Paul, Gaurav Aggarwal, Soma Biswas, Piyush Rai, Kai Han, Vineeth N Balasubramanian*
 - SAFA: Sample-Adaptive Feature Augmentation for Long-Tailed Image Classification, *Yan Hong, Jianfu Zhang, Zhongyi Sun, Ke Yan*
 - Discrete-Constrained Regression for Local Counting Models, *Haipeng Xiong, Angela Yao*
 - Breadcrumbs: Adversarial Class-Balanced Sampling for Long-Tailed Recognition, *Bo Liu, Haoxiang Li, Hao Kang, Gang Hua, Nuno Vasconcelos*
 - Chairs Can Be Stood On: Overcoming Object Bias in Human-Object Interaction Detection, *Guangzhi Wang, Yangyang Guo, Yongkang Wong, Mohan Kankanhalli*
 - DICE: Leveraging Sparsification for Out-of-Distribution Detection, *Yiyu Sun, Yixuan Li*
 - Invariant Feature Learning for Generalized Long-Tailed Classification, *Kaihua Tang, Mingyuan Tao, Jiaxin Qi, Zhengguang Liu, Hanwang Zhang*
 - Centrality and Consistency: Two-Stage Clean Samples Identification for Learning with Instance-Dependent Noisy Labels, *Ganlong Zhao, Guanbin Li, Yipeng Qin, Feng Liu, Yizhou Yu*
 - Hyperspherical Learning in Multi-Label Classification, *Bo Ke, Yunquan Zhu, Mengtian Li, Xiujun Shu, Ruizhi Qiao, Bo Ren*
 - When Active Learning Meets Implicit Semantic Data Augmentation, *Zhuangzhuang Chen, Jin Zhang, Pan Wang, Jie Chen, Jianqiang Li*
 - VL-LTR: Learning Class-Wise Visual-Linguistic Representation for Long-Tailed Visual Recognition, *Changyao Tian, Wenhai Wang, Xizhou Zhu, Jifeng Dai, Yu Qiao*
 - Class Is Invariant to Context and Vice Versa: On Learning Invariance for Out-of-Distribution Generalization, *Jiaxin Qi, Kaihua Tang, Qianru Sun, Xian-Sheng Hua, Hanwang Zhang*
 - Hierarchical Semi-Supervised Contrastive Learning for Contamination-Resistant Anomaly Detection, *Gaoang Wang, Yibing Zhan, Xinchao Wang, Mingli Song, Klara Nahrstedt*
 - Tracking by Associating Clips, *Sanghyun Woo, Kwanyong Park, Seoung Wug Oh, In So Kweon, Joon-Young Lee*
 - Background-Insensitive Scene Text Recognition with Text Semantic Segmentation, *Liang Zhao, Zhenyao Wu, Xinyi Wu, Greg Wilsbacher, Song Wang*
 - Improving Closed and Open-Vocabulary Attribute Prediction Using Transformers, *Khoi Pham, Kushal Kafle, Zhe Lin, Zhihong Ding, Scott Cohen, Quan Tran, Abhinav Shrivastava*
 - Training Vision Transformers with Only 2040 Images, *Yun-Hao Cao, Hao Yu, Jianxin Wu*
 - Bridging Images and Videos: A Simple Learning Framework for Large Vocabulary Video Object Detection, *Sanghyun Woo, Kwanyong Park, Seoung Wug Oh, In So Kweon, Joon-Young Lee*
 - Automatic Check-Out via Prototype-Based Classifier Learning from Single-Product Exemplars, *Hao Chen, Xiu-Shen Wei, Faen Zhang, Yang Shen, Hui Xu, Liang Xiao*
 - Wave-ViT: Unifying Wavelet and Transformers for Visual Representation Learning, *Ting Yao, Yingwei Pan, Yehao Li, Chong-Wah Ngo, Tao Mei*
 - FOSTER: Feature Boosting and Compression for Class-Incremental Learning, *Fu-Yun Wang, Da-Wei Zhou, Han-Jia Ye, De-Chuan Zhan*
 - Improving Fine-Grained Visual Recognition in Low Data Regimes via Self-Boosting Attention Mechanism, *Yangyang Shu, Baosheng Yu, Haiming Xu, Lingqiao Liu*
 - VSA: Learning Varied-Size Window Attention in Vision Transformers, *Qiming Zhang, Yufei Xu, Jing Zhang, Dacheng Tao*
 - Unbiased Manifold Augmentation for Coarse Class Subdivision, *Baoming Yan, Ke Gao, Bo Gao, Lin Wang, Jiang Yang, Xiaobo Li*
 - Rethinking Confidence Calibration for Failure Prediction, *Fei Zhu, Zhen Cheng, Xu-Yao Zhang, Cheng-Lin Liu*
 - Should All Proposals Be Treated Equally in Object Detection? *Yunsheng Li, Yinpeng Chen, Xiyang Dai, Dongdong Chen, Mengchen Liu, Pei Yu, Ying Jin, Lu Yuan, Zicheng Liu, Nuno Vasconcelos*
 - VIP: Unified Certified Detection and Recovery for Patch Attack with Vision Transformers, *Junbo Li, Huan Zhang, Cihang Xie*
 - IGFormer: Interaction Graph Transformer for Skeleton-Based Human Interaction Recognition, *Yunsheng Pang, Qiuhong Ke, Hossein Rahmani, James Bailey, Jun Liu*
 - PRIME: A Few Primitives Can Boost Robustness to Common Corruptions, *Apostolos Modas, Rahul Rade, Guillermo Ortiz-Jiménez, Seyed-Mohsen Moosavi-Dezfooli, Pascal Frossard*
 - Rotation Regularization without Rotation, *Takumi Kobayashi*
 - In Defense of Image Pre-training for Spatiotemporal Recognition, *Xianhang Li, Huiyu Wang, Chen Wei, Jieru Mei, Alan Yuille, Yuyin Zhou, Cihang Xie*
 - Learning with Noisy Labels by Efficient Transition Matrix Estimation to Combat Label Miscalibration, *Seong Min Kye, Kwanghee Choi, Joonyoung Yi, Buru Chang*
 - Online Task-Free Continual Learning with Dynamic Sparse Distributed Memory, *Julien Pourcel, Ngoc-Son Vu, Robert M. French*
- Representation Learning**
- Discriminability-Transferability Trade-Off: An Information-Theoretic Perspective, *Quan Cui, Bingchen Zhao, Zhao-Min Chen, Borui Zhao, Renjie Song, Boyan Zhou, Jiajun Liang, Osamu Yoshie*
 - LocVTP: Video-Text Pre-training for Temporal Localization, *Meng Cao, Tianyu Yang, Junwu Weng, Can Zhang, Jue Wang, Yuexian Zou*
 - Few-Shot End-to-End Object Detection via Constantly Concentrated Encoding across Heads, *Jiawei Ma, Guangxing Han, Shiyuan Huang, Yuncong Yang, Shih-Fu Chang*
 - LiP-Flow: Learning Inference-Time Priors for Codec Avatars via Normalizing Flows in Latent Space, *Emre Aksan, Shugao Ma, Akin Caliskan, Stanislav Pidhorskyi, Alexander Richard, Shih-En Wei, Jason Saragih, Otmar Hilliges*
 - Learning Ego 3D Representation As Ray Tracing, *Jiachen Lu, Zheyuan Zhou, Xiatian Zhu, Hang Xu, Li Zhang*
 - SphereFed: Hyperspherical Federated Learning, *Xin Dong, Sai Qian Zhang, Ang Li, H.T. Kung*
 - Hierarchically Self-Supervised Transformer for Human Skeleton Representation Learning, *Yuxiao Chen, Long Zhao, Jianbo Yuan,*

Yu Tian, Zhaoyang Xia, Shijie Geng, Ligong Han, Dimitris N. Metaxas

- Posterior Refinement on Metric Matrix Improves Generalization Bound in Metric Learning, *Mingda Wang, Canqian Yang, Yi Xu*
- CoSCL: Cooperation of Small Continual Learners Is Stronger than a Big One, *Liyuan Wang, Xingxing Zhang, Qian Li, Jun Zhu, Yi Zhong*
- Fast-MoCo: Boost Momentum-Based Contrastive Learning with Combinatorial Patches, *Yuanzheng Ci, Chen Lin, Lei Bai, Wanli Ouyang*
- LoRD: Local 4D Implicit Representation for High-Fidelity Dynamic Human Modeling, *Boyan Jiang, Xinlin Ren, Mingsong Dou, Xiangyang Xue, Yanwei Fu, Yinda Zhang*
- On the Versatile Uses of Partial Distance Correlation in Deep Learning, *Xingjian Zhen, Zihang Meng, Rudrasis Chakraborty, Vikas Singh*
- Self-Regulated Feature Learning via Teacher-Free Feature Distillation, *Lujun Li*
- Balancing between Forgetting and Acquisition in Incremental Subpopulation Learning, *Mingfu Liang, Jiahuan Zhou, Wei Wei, Ying Wu*
- Counterfactual Intervention Feature Transfer for Visible-Infrared Person Re-identification, *Xulin Li, Yan Lu, Bin Liu, Yating Liu, Guojun Yin, Qi Chu, Jinyang Huang, Feng Zhu, Rui Zhao, Nenghai Yu*
- DAS: Densely-Anchored Sampling for Deep Metric Learning, *Lizhao Liu, Shangxin Huang, Zhuangwei Zhuang, Ran Yang, Minghui Tan, Yaowei Wang*
- Learn from All: Erasing Attention Consistency for Noisy Label Facial Expression Recognition, *Yuhang Zhang, Chengrui Wang, Xu Ling, Weihong Deng*
- UFO: Unified Feature Optimization, *Teng Xi, Yifan Sun, Deli Yu, Bi Li, Nan Peng, Gang Zhang, Xinyu Zhang, Zhigang Wang, Jinwen Chen, Jian Wang, Lufei Liu, Haocheng Feng, Junyu Han, Jingtuo Liu, Errui Ding, Jingdong Wang*
- X-Learner: Learning Cross Sources and Tasks for Universal Visual Representation, *Yinan He, Gengshi Huang, Siyu Chen, Jianing Teng, Kun Wang, Zhenfei Yin, Lu Sheng, Ziwei Liu, Yu Qiao, Jing Shao*
- ParC-Net: Position Aware Circular Convolution with Merits from ConvNets and Transformer, *Haokui Zhang, Wenze Hu, Xiaoyu Wang*
- DualPrompt: Complementary Prompting for Rehearsal-Free Continual Learning, *Zifeng Wang, Zizhao Zhang, Sayna Ebrahimi, Ruoxi Sun, Han Zhang, Chen-Yu Lee, Xiaoqi Ren, Guolong Su, Vincent Perot, Jennifer Dy, Tomas Pfister*
- Unifying Visual Contrastive Learning for Object Recognition from a Graph Perspective, *Shixiang Tang, Feng Zhu, Lei Bai, Rui Zhao, Chenyu Wang, Wanli Ouyang*
- Conditional Stroke Recovery for Fine-Grained Sketch-Based Image Retrieval, *Zhixin Ling, Zhen Xing, Jian Zhou, Xiangdong Zhou*
- Relative Contrastive Loss for Unsupervised Representation Learning, *Shixiang Tang, Feng Zhu, Lei Bai, Rui Zhao, Wanli Ouyang*
- Fine-Grained Fashion Representation Learning by Online Deep Clustering, *Yang Jiao, Ning Xie, Yan Gao, Chien-chih Wang, Yi Sun*
- A Gyrovector Space Approach for Symmetric Positive Semi-Definite Matrix Learning, *Xuan Son Nguyen*
- Learning Visual Representation from Modality-Shared Contrastive Language-Image Pre-training, *Haoxuan You,*

Luowei Zhou, Bin Xiao, Noel Codella, Yu Cheng, Ruochen Xu, Shih-Fu Chang, Lu Yuan

- Class-Incremental Learning with Cross-Space Clustering and Controlled Transfer, *Arjun Ashok, K J Joseph, Vineeth N Balasubramanian*
 - MVDG: A Unified Multi-View Framework for Domain Generalization, *Jian Zhang, Lei Qi, Yinghuan Shi, Yang Gao*
- Scene Analysis and Understanding**
- RigNet: Repetitive Image Guided Network for Depth Completion, *Zhiqiang Yan, Kun Wang, Xiang Li, Zhenyu Zhang, Jun Li, Jian Yang*
 - FADE: Fusing the Assets of Decoder and Encoder for Task-Agnostic Upsampling, *Hao Lu, Wenze Liu, Hongtao Fu, Zhiguo Cao*
 - Hierarchical Memory Learning for Fine-Grained Scene Graph Generation, *Youming Deng, Yansheng Li, Yongjun Zhang, Xiang Xiang, Jian Wang, Jingdong Chen, Jiayi Ma*
 - DODA: Data-Oriented Sim-to-Real Domain Adaptation for 3D Semantic Segmentation, *Runyu Ding, Jihan Yang, Li Jiang, Xiaojuan Qi*
 - MonoPLFlowNet: Permutohedral Lattice FlowNet for Real-Scale 3D Scene Flow Estimation with Monocular Images, *Runfa Li, Truong Nguyen*
 - Is It Necessary to Transfer Temporal Knowledge for Domain Adaptive Video Semantic Segmentation? *Xinyi Wu, Zhenyao Wu, Jin Wan, Lili Ju, Song Wang*
 - Towards Hard-Positive Query Mining for DETR-Based Human-Object Interaction Detection, *Xubin Zhong, Changxing Ding, Zijian Li, Shaoli Huang*
 - Discovering Human-Object Interaction Concepts via Self-Compositional Learning, *Zhi Hou, Baosheng Yu, Dacheng Tao*
 - Stereo Depth Estimation with Echoes, *Chenghao Zhang, Kun Tian, Bolin Ni, Gaofeng Meng, Bin Fan, Zhaoxiang Zhang, Chunhong Pan*
 - PETR: Position Embedding Transformation for Multi-View 3D Object Detection, *Yingfei Liu, Tiancai Wang, Xiangyu Zhang, Jian Sun*
 - S2Net: Stochastic Sequential Pointcloud Forecasting, *Xinshuo Weng, Junyu Nan, Kuan-Hui Lee, Rowan McAllister, Adrien Gaidon, Nicholas Rhinehart, Kris M. Kitani*
 - RA-Depth: Resolution Adaptive Self-Supervised Monocular Depth Estimation, *Mu He, Le Hui, Yikai Bian, Jian Ren, Jin Xie, Jian Yang*
 - PolyphonicFormer: Unified Query Learning for Depth-Aware Video Panoptic Segmentation, *Haobo Yuan, Xiangtai Li, Yibo Yang, Guangliang Cheng, Jing Zhang, Yunhai Tong, Lefei Zhang, Dacheng Tao*
 - SQN: Weakly-Supervised Semantic Segmentation of Large-Scale 3D Point Clouds, *Qingyong Hu, Bo Yang, Guangchi Fang, Yulan Guo, Aleš Leonardis, Niki Trigoni, Andrew Markham*
 - Initialization and Alignment for Adversarial Texture Optimization, *Xiaoming Zhao, Zhizhen Zhao, Alexander G. Schwing*
 - MOTR: End-to-End Multiple-Object Tracking with TRansformer, *Fangao Zeng, Bin Dong, Yuang Zhang, Tiancai Wang, Xiangyu Zhang, Yichen Wei*
 - GALA: Toward Geometry-and-Lighting-Aware Object Search for Compositing, *Sijie Zhu, Zhe Lin, Scott Cohen, Jason Kuen, Zhifei Zhang, Chen Chen*

- Panoptic-PartFormer: Learning a Unified Model for Panoptic Part Segmentation, *Xiangtai Li, Shilin Xu, Yibo Yang, Guangliang Cheng, Yunhai Tong, Dacheng Tao*
 - Salient Object Detection for Point Clouds, *Songlin Fan, Wei Gao, Ge Li*
 - Weakly Supervised 3D Scene Segmentation with Region-Level Boundary Awareness and Instance Discrimination, *Kangcheng Liu, Yuzhi Zhao, Qiang Nie, Zhi Gao, Ben M. Chen*
 - Towards Open-Vocabulary Scene Graph Generation with Prompt-Based Finetuning, *Tao He, Lianli Gao, Jingkuan Song, Yuan-Fang Li*
 - Variance-Aware Weight Initialization for Point Convolutional Neural Networks, *Pedro Hermosilla, Michael Schelling, Tobias Ritschel, Timo Ropinski*
 - 3DG-STFM: 3D Geometric Guided Student-Teacher Feature Matching, *Runyu Mao, Chen Bai, Yatang An, Fengqing Zhu, Cheng Lu*
 - Video Restoration Framework and Its Meta-Adaptations to Data-Poor Conditions, *Prashant W Patil, Sunil Gupta, Santu Rana, Svetha Venkatesh*
- Scene Text and Document Understanding**
- When Counting Meets HMER: Counting-Aware Network for Handwritten Mathematical Expression Recognition, *Bohan Li, Ye Yuan, Dingkang Liang, Xiao Liu, Zhilong Ji, Jinfeng Bai, Wenyu Liu, Xiang Bai*
 - Detecting Tampered Scene Text in the Wild, *Yuxin Wang, Hongtao Xie, Mengting Xing, Jing Wang, Shenggao Zhu, Yongdong Zhang*
 - Optimal Boxes: Boosting End-to-End Scene Text Recognition by Adjusting Annotated Bounding Boxes via Reinforcement Learning, *Jingqun Tang, Wenming Qian, Luchuan Song, Xiena Dong, Lan Li, Xiang Bai*
 - COO: Comic Onomatopoeia Dataset for Recognizing Arbitrary or Truncated Texts, *Jeonghun Baek, Yusuke Matsui, Kiyoharu Aizawa*
 - Levenshtein OCR, *Cheng Da, Peng Wang, Cong Yao*
 - Multi-Granularity Prediction for Scene Text Recognition, *Peng Wang, Cheng Da, Cong Yao*
 - Dynamic Low-Resolution Distillation for Cost-Efficient End-to-End Text Spotting, *Ying Chen, Liang Qiao, Zhanzhan Cheng, Shiliang Pu, Yi Niu, Xi Li*
 - Contextual Text Block Detection towards Scene Text Understanding, *Chuhui Xue, Jiaxing Huang, Wenqing Zhang, Shijian Lu, Changhu Wang, Song Bai*
 - CoMER: Modeling Coverage for Transformer-Based Handwritten Mathematical Expression Recognition, *Wenqi Zhao, Liangcai Gao*
 - Don't Forget Me: Accurate Background Recovery for Text Removal via Modeling Local-Global Context, *Chongyu Liu, Lianwen Jin, Yuliang Liu, Canjie Luo, Bangdong Chen, Fengjun Guo, Kai Ding*
 - SGBANet: Semantic GAN and Balanced Attention Network for Arbitrarily Oriented Scene Text Recognition, *Dajian Zhong, Shujing Lyu, Palaiahnakote Shivakumara, Bing Yin, Jiajia Wu, Umapada Pal, Yue Lu*
- Segmentation, Grouping, and Shape**
- CAR: Class-Aware Regularizations for Semantic Segmentation, *Ye Huang, Di Kang, Liang Chen, Xuefei Zhe, Wenjing Jia, Linchao Bao, Xiangjian He*
 - Saliency Hierarchy Modeling via Generative Kernels for Salient Object Detection, *Wenhu Zhang, Liangli Zheng, Huanyu Wang, Xintian Wu, Xi Li*
 - Active Pointly-Supervised Instance Segmentation, *Chufeng Tang, Lingxi Xie, Gang Zhang, Xiaopeng Zhang, Qi Tian, Xiaolin Hu*
 - A Transformer-Based Decoder for Semantic Segmentation with Multi-level Context Mining, *Bowen Shi, Dongsheng Jiang, Xiaopeng Zhang, Han Li, Wenrui Dai, Junni Zou, Hongkai Xiong, Qi Tian*
 - XMem: Long-Term Video Object Segmentation with an Atkinson-Shiffrin Memory Model, *Ho Kei Cheng, Alexander G. Schwing*
 - Self-Distillation for Robust LiDAR Semantic Segmentation in Autonomous Driving, *Jiale Li, Hang Dai, Yong Ding*
 - 2DPASS: 2D Priors Assisted Semantic Segmentation on LiDAR Point Clouds, *Xu Yan, Jiantao Gao, Chaoda Zheng, Chao Zheng, Ruimao Zhang, Shuguang Cui, Zhen Li*
 - Box-Supervised Instance Segmentation with Level Set Evolution, *Wentong Li, Wenyu Liu, Jianke Zhu, Miaomiao Cui, Xian-Sheng Hua, Lei Zhang*
 - Point Primitive Transformer for Long-Term 4D Point Cloud Video Understanding, *Hao Wen, Yunze Liu, Jingwei Huang, Bo Duan, Li Yi*
 - Adaptive Agent Transformer for Few-Shot Segmentation, *Yuan Wang, Rui Sun, Zhe Zhang, Tianzhu Zhang*
 - Waymo Open Dataset: Panoramic Video Panoptic Segmentation, *Jieru Mei, Alex Zihao Zhu, Xinchun Yan, Hang Yan, Siyuan Qiao, Yukun Zhu, Liang-Chieh Chen, Henrik Kretschmar*
 - AdaAfford: Learning to Adapt Manipulation Affordance for 3D Articulated Objects via Few-Shot Interactions, *Yian Wang, Ruihai Wu, Kaichun Mo, Jiaqi Ke, Qingnan Fan, Leonidas J. Guibas, Hao Dong*
 - 2D Amodal Instance Segmentation Guided by 3D Shape Prior, *Zhixuan Li, Weining Ye, Tingting Jiang, Tiejun Huang*
 - Data Efficient 3D Learner via Knowledge Transferred from 2D Model, *Ping-Chung Yu, Cheng Sun, Min Sun*
 - Adaptive Spatial-BCE Loss for Weakly Supervised Semantic Segmentation, *Tong Wu, Guangyu Gao, Junshi Huang, Xiaolin Wei, Xiaoming Wei, Chi Harold Liu*
 - 3D Instances as 1D Kernels, *Yizheng Wu, Min Shi, Shuaiyuan Du, Hao Lu, Zhiguo Cao, Weicai Zhong*
 - TransMatting: Enhancing Transparent Objects Matting with Transformers, *Huanqia Cai, Fanglei Xue, Lele Xu, Lili Guo*
 - MVSaNet: Multi-View Augmentation for RGB-D Salient Object Detection, *Jiayuan Zhou, Lijun Wang, Huchuan Lu, Kaining Huang, Xinchu Shi, Bocong Liu*
 - k-Means Mask Transformer, *Qihang Yu, Huiyu Wang, Siyuan Qiao, Maxwell Collins, Yukun Zhu, Hartwig Adam, Alan Yuille, Liang-Chieh Chen*
 - SegPGD: An Effective and Efficient Adversarial Attack for Evaluating and Boosting Segmentation Robustness, *Jindong Gu, Hengshuang Zhao, Volker Tresp, Philip H. S. Torr*
 - Adversarial Erasing Framework via Triplet with Gated Pyramid Pooling Layer for Weakly Supervised Semantic Segmentation, *Sung-Hoon Yoon, Hyeokjun Kweon, Jegyeong Cho, Shinjeong Kim, Kuk-Jin Yoon*
 - Continual Semantic Segmentation via Structure Preserving and Projected Feature Alignment, *Zihan Lin, Zilei Wang, Yixin Zhang*
 - Interclass Prototype Relation for Few-Shot Segmentation, *Atsuro Okazawa*

- Slim Scissors: Segmenting Thin Object from Synthetic Background, *Kunyang Han, Jun Hao Liew, Jiashi Feng, Huawei Tian, Yao Zhao, Yunchao Wei*
 - Learning Quality-Aware Dynamic Memory for Video Object Segmentation, *Yong Liu, Ran Yu, Fei Yin, Xinyuan Zhao, Wei Zhao, Weihao Xia, Yujiu Yang*
 - Learning Implicit Feature Alignment Function for Semantic Segmentation, *Hanzhe Hu, Yinbo Chen, Jiarui Xu, Shubhankar Borse, Hong Cai, Fatih Porikli, Xiaolong Wang*
 - Instance As Identity: A Generic Online Paradigm for Video Instance Segmentation, *Feng Zhu, Zongxin Yang, Xin Yu, Yi Yang, Yunchao Wei*
 - Laplacian Mesh Transformer: Dual Attention and Topology Aware Network for 3D Mesh Classification and Segmentation, *Xiao-Juan Li, Jie Yang, Fang-Lue Zhang*
 - Geodesic-Former: A Geodesic-Guided Few-Shot 3D Point Cloud Instance Segmenter, *Tuan Ngo, Khoi Nguyen*
 - Union-Set Multi-source Model Adaptation for Semantic Segmentation, *Zongyao Li, Ren Togo, Takahiro Ogawa, Miki Haseyama*
 - Global Spectral Filter Memory Network for Video Object Segmentation, *Yong Liu, Ran Yu, Jiahao Wang, Xinyuan Zhao, Yitong Wang, Yansong Tang, Yujiu Yang*
 - A Simple Baseline for Open-Vocabulary Semantic Segmentation with Pre-trained Vision-Language Model, *Mengde Xu, Zheng Zhang, Fangyun Wei, Yutong Lin, Yue Cao, Han Hu, Xiang Bai*
 - UCTNet: Uncertainty-Aware Cross-Modal Transformer Network for Indoor RGB-D Semantic Segmentation, *Xiaowen Ying, Mooi Choo Chuah*
 - Learning Regional Purity for Instance Segmentation on 3D Point Clouds, *Shichao Dong, Guosheng Lin, Tzu-Yi Hung*
 - Cross-Domain Few-Shot Semantic Segmentation, *Shuo Lei, Xuchao Zhang, Jianfeng He, Fanglan Chen, Bowen Du, Chang-Tien Lu*
- Semi / Weak / Self / Unsupervised Learning**
- Generative Subgraph Contrast for Self-Supervised Graph Representation Learning, *Yuehui Han, Le Hui, Haobo Jiang, Jianjun Qian, Jin Xie*
 - SdAE: Self-Distilled Masked Autoencoder, *Yabo Chen, Yuchen Liu, Dongsheng Jiang, Xiaopeng Zhang, Wenrui Dai, Hongkai Xiong, Qi Tian*
 - Vibration-Based Uncertainty Estimation for Learning from Limited Supervision, *Hengtong Hu, Lingxi Xie, Xinyue Huo, Richang Hong, Qi Tian*
 - Weakly Supervised Object Localization through Inter-class Feature Similarity and Intra-Class Appearance Consistency, *Jun Wei, Sheng Wang, S. Kevin Zhou, Shuguang Cui, Zhen Li*
 - Mc-BEiT: Multi-Choice Discretization for Image BERT Pre-training, *Xiaotong Li, Yixiao Ge, Kun Yi, Zixuan Hu, Ying Shan, Ling-Yu Duan*
 - Bootstrapped Masked Autoencoders for Vision BERT Pretraining, *Xiaoyi Dong, Jianmin Bao, Ting Zhang, Dongdong Chen, Weiming Zhang, Lu Yuan, Dong Chen, Fang Wen, Nenghai Yu*
 - Unsupervised Visual Representation Learning by Synchronous Momentum Grouping, *Bo Pang, Yifan Zhang, Yaoyi Li, Jia Cai, Cewu Lu*
 - Improving Few-Shot Part Segmentation Using Coarse Supervision, *Oindrila Saha, Zezhou Cheng, Subhansu Maji*
 - MVP: Multimodality-Guided Visual Pre-training, *Longhui Wei, Lingxi Xie, Wengang Zhou, Houqiang Li, Qi Tian*
 - Unsupervised Selective Labeling for More Effective Semi-Supervised Learning, *Xudong Wang, Long Lian, Stella X. Yu*
 - Dense Siamese Network for Dense Unsupervised Learning, *Wenwei Zhang, Jiangmiao Pang, Kai Chen, Chen Change Loy*
 - Multi-Granularity Distillation Scheme towards Lightweight Semi-Supervised Semantic Segmentation, *Jie Qin, Jie Wu, Ming Li, Xuefeng Xiao, Min Zheng, Xingang Wang*
 - CP2: Copy-Paste Contrastive Pretraining for Semantic Segmentation, *Feng Wang, Huiyu Wang, Chen Wei, Alan Yuille, Wei Shen*
 - Self-Filtering: A Noise-Aware Sample Selection for Label Noise with Confidence Penalization, *Qi Wei, Haoliang Sun, Xiankai Lu, Yilong Yin*
 - RDA: Reciprocal Distribution Alignment for Robust Semi-Supervised Learning, *Yue Duan, Lei Qi, Lei Wang, Luping Zhou, Yinghuan Shi*
 - United Defocus Blur Detection and Deblurring via Adversarial Promoting Learning, *Wenda Zhao, Fei Wei, You He, Huchuan Lu*
 - Semi-Supervised Vision Transformers, *Zejia Weng, Xitong Yang, Ang Li, Zuxuan Wu, Yu-Gang Jiang*
 - Diverse Learner: Exploring Diverse Supervision for Semi-Supervised Object Detection, *Linfeng Li, Minyue Jiang, Yue Yu, Wei Zhang, Xiangru Lin, Yingying Li, Xiao Tan, Jingdong Wang, Errui Ding*
 - FedX: Unsupervised Federated Learning with Cross Knowledge Distillation, *Sungwon Han, Sungwon Park, Fangzhao Wu, Sundong Kim, Chuhan Wu, Xing Xie, Meeyoung Cha*
 - W2N: Switching from Weak Supervision to Noisy Supervision for Object Detection, *Zitong Huang, Yiping Bao, Bowen Dong, Erjin Zhou, Wangmeng Zuo*
 - GOCA: Guided Online Cluster Assignment for Self-Supervised Video Representation Learning, *Huseyin Coskun, Alireza Zareian, Joshua L. Moore, Federico Tombari, Chen Wang*
 - Revisiting the Critical Factors of Augmentation-Invariant Representation Learning, *Junqiang Huang, Xiangwen Kong, Xiangyu Zhang*
 - CA-SSL: Class-Agnostic Semi-Supervised Learning for Detection and Segmentation, *Lu Qi, Jason Kuen, Zhe Lin, Jiuxiang Gu, Fengyun Rao, Dian Li, Weidong Guo, Zhen Wen, Ming-Hsuan Yang, Jiaya Jia*
 - Domain Invariant Masked Autoencoders for Self-Supervised Learning from Multi-Domains, *Haiyang Yang, Shixiang Tang, Meilin Chen, Yizhou Wang, Feng Zhu, Lei Bai, Rui Zhao, Wanli Ouyang*
 - Coarse-to-Fine Incremental Few-Shot Learning, *Xiang Xiang, Yuwen Tan, Qian Wan, Jing Ma, Alan Yuille, Gregory D. Hager*
 - Learning Unbiased Transferability for Domain Adaptation by Uncertainty Modeling, *Jian Hu, Haowen Zhong, Fei Yang, Shaogang Gong, Guile Wu, Junchi Yan*
 - Stochastic Consensus: Enhancing Semi-Supervised Learning with Consistency of Stochastic Classifiers, *Hui Tang, Lin Sun, Kui Jia*
 - DiffuseMorph: Unsupervised Deformable Image Registration Using Diffusion Model, *Boah Kim, Inhwa Han, Jong Chul Ye*
 - Semi-Leak: Membership Inference Attacks against Semi-Supervised Learning, *Xinlei He, Hongbin Liu, Neil Zhenqiang Gong, Yang Zhang*
 - OpenLDN: Learning to Discover Novel Classes for Open-World Semi-Supervised Learning, *Mamshad Nayeem Rizve, Navid Kardan, Salman Khan, Fahad Shahbaz Khan, Mubarak Shah*

- Unsupervised Few-Shot Image Classification by Learning Features into Clustering Space, *Shuo Li, Fang Liu, Zehua Hao, Kaibo Zhao, Licheng Jiao*

- PseudoAugment: Learning to Use Unlabeled Data for Data Augmentation in Point Clouds, *Zhaoqi Leng, Shuyang Cheng, Benjamin Caine, Weiyue Wang, Xiao Zhang, Jonathon Shlens, Mingxing Tan, Dragomir Anguelov*

Stereo, 3D from Multiview and Other Sensors

- MVSTER: Epipolar Transformer for Efficient Multi-View Stereo, *Xiaofeng Wang, Zheng Zhu, Guan Huang, Fangbo Qin, Yun Ye, Yijia He, Xu Chi, Xingang Wang*
- KD-MVS: Knowledge Distillation Based Self-Supervised Learning for Multi-View Stereo, *Yikang Ding, Qingtian Zhu, Xiangyue Liu, Wentao Yuan, Haotian Zhang, Chi Zhang*
- SALVe: Semantic Alignment Verification for Floorplan Reconstruction from Sparse Panoramas, *John Lambert, Yuguang Li, Iyaylo Boyadzhiev, Lambert Wixson, Manjunath Narayana, Will Hutchcroft, James Hays, Frank Dellaert, Sing Bing Kang*
- NDF: Neural Deformable Fields for Dynamic Human Modelling, *Ruiqi Zhang, Jie Chen*
- NeXT: Towards High Quality Neural Radiance Fields via Multi-Skip Transformer, *Yunxiao Wang, Yanjie Li, Peidong Liu, Tao Dai, Shu-Tao Xia*
- Decomposing the Tangent of Occluding Boundaries according to Curvatures and Torsions, *Huizong Yang, Anthony Yezzi*
- NeuRIS: Neural Reconstruction of Indoor Scenes Using Normal Priors, *Jiepeng Wang, Peng Wang, Xiaoxiao Long, Christian Theobalt, Taku Komura, Lingjie Liu, Wenping Wang*
- Improving RGB-D Point Cloud Registration by Learning Multi-Scale Local Linear Transformation, *Ziming Wang, Xiaoliang Huo, Zhenghao Chen, Jing Zhang, Lu Sheng, Dong Xu*
- Real-Time Neural Character Rendering with Pose-Guided Multiplane Images, *Hao Ouyang, Bo Zhang, Pan Zhang, Hao Yang, Jiaolong Yang, Dong Chen, Qifeng Chen, Fang Wen*
- SparseNeuS: Fast Generalizable Neural Surface Reconstruction from Sparse Views, *Xiaoxiao Long, Cheng Lin, Peng Wang, Taku Komura, Wenping Wang*
- Disentangling Object Motion and Occlusion for Unsupervised Multi-Frame Monocular Depth, *Ziyue Feng, Liang Yang, Longlong Jing, Haiyan Wang, YingLi Tian, Bing Li*
- Gen6D: Generalizable Model-Free 6-DoF Object Pose Estimation from RGB Images, *Yuan Liu, Yilin Wen, Sida Peng, Cheng Lin, Xiaoxiao Long, Taku Komura, Wenping Wang*
- Latency-Aware Collaborative Perception, *Zixing Lei, Shunli Ren, Yue Hu, Wenjun Zhang, Siheng Chen*
- TensorRF: Tensorial Radiance Fields, *Anpei Chen, Zexiang Xu, Andreas Geiger, Jingyi Yu, Hao Su*
- EASNet: Searching Elastic and Accurate Network Architecture for Stereo Matching, *Qiang Wang, Shaohuai Shi, Kaiyong Zhao, Xiaowen Chu*
- Selection and Cross Similarity for Event-Image Deep Stereo, *Hoonhee Cho, Kuk-Jin Yoon*
- CIRCLE: Convolutional Implicit Reconstruction and Completion for Large-Scale Indoor Scene, *Hao-Xiang Chen, Jiahui Huang, Tai-Jiang Mu, Shi-Min Hu*
- ParticleSfM: Exploiting Dense Point Trajectories for Localizing Moving Cameras in the Wild, *Wang Zhao, Shaohui Liu, Hengkai Guo, Wenping Wang, Yong-Jin Liu*
- CoVisPose: Co-Visibility Pose Transformer for Wide-Baseline Relative Pose Estimation in 360° Indoor Panoramas, *Will*

Hutchcroft, Yuguang Li, Iyaylo Boyadzhiev, Zhiqiang Wan, Haiyan Wang, Sing Bing Kang

- GraphFit: Learning Multi-Scale Graph-Convolutional Representation for Point Cloud Normal Estimation, *Keqiang Li, Mingyang Zhao, Huaiyu Wu, Dong-Ming Yan, Zhen Shen, Fei-Yue Wang, Gang Xiong*
- Point Scene Understanding via Disentangled Instance Mesh Reconstruction, *Jiaxiang Tang, Xiaokang Chen, Jingbo Wang, Gang Zeng*
- Structure and Motion from Casual Videos, *Zhoutong Zhang, Forrester Cole, Zhengqi Li, Noah Snavely, Michael Rubinstein, William T. Freeman*
- GraphCSPN: Geometry-Aware Depth Completion via Dynamic GCNs, *Xin Liu, Xiaofei Shao, Bo Wang, Yali Li, Shengjin Wang*
- MODE: Multi-View Omnidirectional Depth Estimation with 360° Cameras, *Ming Li, Xueqian Jin, Xuejiao Hu, Jingzhao Dai, Sidan Du, Yang Li*
- ActiveNeRF: Learning Where to See with Uncertainty Estimation, *Xuran Pan, Zihang Lai, Shiji Song, Gao Huang*

Transfer Learning

- Class-Incremental Novel Class Discovery, *Subhankar Roy, Mingxuan Liu, Zhun Zhong, Nicu Sebe, Elisa Ricci*
- Unknown-Oriented Learning for Open Set Domain Adaptation, *Jie Liu, Xiaoqing Guo, Yixuan Yuan*
- Prototype-Guided Continual Adaptation for Class-Incremental Unsupervised Domain Adaptation, *Hongbin Lin, Yifan Zhang, Zhen Qiu, Shuaicheng Niu, Chuang Gan, Yanxia Liu, Mingkui Tan*
- DecoupleNet: Decoupled Network for Domain Adaptive Semantic Segmentation, *Xin Lai, Zhuotao Tian, Xiaogang Xu, Yingcong Chen, Shu Liu, Hengshuang Zhao, Liwei Wang, Jiaya Jia*
- Class-Agnostic Object Counting Robust to Intraclass Diversity, *Shenjian Gong, Shanshan Zhang, Jian Yang, Dengxin Dai, Bernt Schiele*
- Burn after Reading: Online Adaptation for Cross-Domain Streaming Data, *Luyu Yang, Mingfei Gao, Zeyuan Chen, Ran Xu, Abhinav Shrivastava, Chetan Ramaiah*
- Mind the Gap in Distilling StyleGANs, *Guodong Xu, Yuenan Hou, Ziwei Liu, Chen Change Loy*
- RegionCL: Exploring Contrastive Region Pairs for Self-Supervised Representation Learning, *Yufei Xu, Qiming Zhang, Jing Zhang, Dacheng Tao*
- Long-Tailed Class Incremental Learning, *Xialei Liu, Yu-Song Hu, Xu-Sheng Cao, Andrew D. Bagdanov, Ke Li, Ming-Ming Cheng*
- Adversarial Partial Domain Adaptation by Cycle Inconsistency, *Kun-Yu Lin, Jiaming Zhou, Yukun Qiu, Wei-Shi Zheng*
- Prior Knowledge Guided Unsupervised Domain Adaptation, *Tao Sun, Cheng Lu, Haibin Ling*
- AcroFOD: An Adaptive Method for Cross-Domain Few-Shot Object Detection, *Yipeng Gao, Lingxiao Yang, Yunmu Huang, Song Xie, Shiyong Li, Wei-Shi Zheng*
- Quasi-Balanced Self-Training on Noise-Aware Synthesis of Object Point Clouds for Closing Domain Gap, *Yongwei Chen, Zihao Wang, Longkun Zou, Ke Chen, Kui Jia*
- Interpretable Open-Set Domain Adaptation via Angular Margin Separation, *Xinhao Li, Jingjing Li, Zhekai Du, Lei Zhu, Wen Li*
- Prototypical Contrast Adaptation for Domain Adaptive Semantic Segmentation, *Zhengkai Jiang, Yuxi Li, Ceyuan Yang, Peng Gao, Yabiao Wang, Ying Tai, Chengjie Wang*
- RBC: Rectifying the Biased Context in Continual Semantic Segmentation, *Hanbin Zhao, Fengyu Yang, Xinghe Fu, Xi Li*

- Factorizing Knowledge in Neural Networks, *Xingyi Yang, Jingwen Ye, Xinchao Wang*
- BMD: A General Class-Balanced Multicentric Dynamic Prototype Strategy for Source-Free Domain Adaptation, *Sanqing Qu, Guang Chen, Jing Zhang, Zhijun Li, Wei He, Dacheng Tao*
- Generalized Brain Image Synthesis with Transferable Convolutional Sparse Coding Networks, *Yawen Huang, Feng Zheng, Xu Sun, Yuexiang Li, Ling Shao, Yefeng Zheng*
- Incomplete Multi-View Domain Adaptation via Channel Enhancement and Knowledge Transfer, *Haifeng Xia, Pu Wang, Zhengming Ding*
- DistPro: Searching a Fast Knowledge Distillation Process via Meta Optimization, *Xueqing Deng, Dawei Sun, Shawn Newsam, Peng Wang*
- ML-BPM: Multi-Teacher Learning with Bidirectional Photometric Mixing for Open Compound Domain Adaptation in Semantic Segmentation, *Fei Pan, Sungsu Hur, Seokju Lee, Junsik Kim, In So Kweon*
- Personalized Education: Blind Knowledge Distillation, *Xiang Deng, Jian Zheng, Zhongfei Zhang*
- Not All Models Are Equal: Predicting Model Transferability in a Self-Challenging Fisher Space, *Wenqi Shao, Xun Zhao, Yixiao Ge, Zhaoyang Zhang, Lei Yang, Xiaogang Wang, Ying Shan, Ping Luo*
- Attention Diversification for Domain Generalization, *Rang Meng, Xianfeng Li, Weijie Chen, Shicai Yang, Jie Song, Xinchao Wang, Lei Zhang, Mingli Song, Di Xie, Shiliang Pu*
- ESS: Learning Event-Based Semantic Segmentation from Still Images, *Zhaoning Sun, Nico Messikommer, Daniel Gehrig, Davide Scaramuzza*

Video Analysis and Understanding

- An Efficient Spatio-Temporal Pyramid Transformer for Action Detection, *Yuetian Weng, Zizheng Pan, Mingfei Han, Xiaojun Chang, Bohan Zhuang*
- Human Trajectory Prediction via Neural Social Physics, *Jiangbei Yue, Dinesh Manocha, He Wang*
- Towards Open Set Video Anomaly Detection, *Yuansheng Zhu, Wentao Bao, Qi Yu*
- Joint-Modal Label Denoising for Weakly-Supervised Audio-Visual Video Parsing, *Haoyue Cheng, Zhaoyang Liu, Hang Zhou, Chen Qian, Wayne Wu, Limin Wang*
- Adaptive Face Forgery Detection in Cross Domain, *Luchuan Song, Zheng Fang, Xiaodan Li, Xiaoyi Dong, Zhenchao Jin, Yuefeng Chen, Siwei Lyu*
- TALLFormer: Temporal Action Localization with a Long-Memory Transformer, *Feng Cheng, Gedas Bertasius*
- Mining Relations among Cross-Frame Affinities for Video Semantic Segmentation, *Guolei Sun, Yun Liu, Hao Tang, Ajad Chhatkuli, Le Zhang, Luc Van Gool*
- Rethinking Learning Approaches for Long-Term Action Anticipation, *Megha Nawhal, Akash Abdu Jyothi, Greg Mori*
- Hierarchical Feature Alignment Network for Unsupervised Video Object Segmentation, *Gensheng Pei, Fumin Shen, Yazhou Yao, Guo-Sen Xie, Zhenmin Tang, Jinhui Tang*
- PAC-Net: Highlight Your Video via History Preference Modeling, *Hang Wang, Penghao Zhou, Chong Zhou, Zhao Zhang, Xing Sun*
- A Sliding Window Scheme for Online Temporal Action Localization, *Young Hwi Kim, Hyolim Kang, Seon Joo Kim*

- ERA: Expert Retrieval and Assembly for Early Action Prediction, *Lin Geng Foo, Tianjiao Li, Hossein Rahmani, Qihong Ke, Jun Liu*
- NSNet: Non-Saliency Suppression Sampler for Efficient Video Recognition, *Boyang Xia, Wenhao Wu, Haoran Wang, Rui Su, Dongliang He, Haosen Yang, Xiaoran Fan, Wanli Ouyang*
- Video Activity Localisation with Uncertainties in Temporal Boundary, *Jiabo Huang, Hailin Jin, Shaogang Gong, Yang Liu*
- Temporal Saliency Query Network for Efficient Video Recognition, *Boyang Xia, Zhihao Wang, Wenhao Wu, Haoran Wang, Jungong Han*
- Efficient Video Transformers with Spatial-Temporal Token Selection, *Junke Wang, Xitong Yang, Hengduo Li, Li Liu, Zuxuan Wu, Yu-Gang Jiang*
- Long Movie Clip Classification with State-Space Video Models, *Md Mohaiminul Islam, Gedas Bertasius*
- Prompting Visual-Language Models for Efficient Video Understanding, *Chen Ju, Tengda Han, Kunhao Zheng, Ya Zhang, Weidi Xie*
- Asymmetric Relation Consistency Reasoning for Video Relation Grounding, *Huan Li, Ping Wei, Jiapeng Li, Zeyu Ma, Jiahui Shang, Nanning Zheng*
- MorphMLP: An Efficient MLP-Like Backbone for Spatial-Temporal Representation Learning, *David Junhao Zhang, Kunchang Li, Yali Wang, Yunpeng Chen, Shashwat Chandra, Yu Qiao, Luoqi Liu, Mike Zheng Shou*
- E-NeRV: Expedite Neural Video Representation with Disentangled Spatial-Temporal Context, *Zizhang Li, Mengmeng Wang, Huaijin Pi, Kechun Xu, Jianbiao Mei, Yong Liu*
- Deep 360° Optical Flow Estimation Based on Multi-Projection Fusion, *Yiheng Li, Connelly Barnes, Kun Huang, Fang-Lue Zhang*
- Learning Long-Term Spatial-Temporal Graphs for Active Speaker Detection, *Kyle Min, Sourya Roy, Subarna Tripathi, Tanaya Guha, Somdeb Majumdar*
- Motion Sensitive Contrastive Learning for Self-Supervised Video Representation, *Jingcheng Ni, Nan Zhou, Jie Qin, Qian Wu, Junqi Liu, Boxun Li, Di Huang*
- Dynamic Temporal Filtering In Video Models, *Fuchen Long, Zhaofan Qiu, Yingwei Pan, Ting Yao, Chong-Wah Ngo, Tao Mei*

Vision + Language

- Tip-Adapter: Training-Free Adaption of CLIP for Few-Shot Classification, *Renrui Zhang, Wei Zhang, Rongyao Fang, Peng Gao, Kunchang Li, Jifeng Dai, Yu Qiao, Hongsheng Li*
- Temporal Lift Pooling for Continuous Sign Language Recognition, *Lianyu Hu, Liqing Gao, Zekang Liu, Wei Feng*
- MORE: Multi-Order RELation Mining for Dense Captioning in 3D Scenes, *Yang Jiao, Shaoxiang Chen, Zequn Jie, Jingjing Chen, Lin Ma, Yu-Gang Jiang*
- SiRi: A Simple Selective Retraining Mechanism for Transformer-Based Visual Grounding, *Mengxue Qu, Yu Wu, Wu Liu, Qiqi Gong, Xiaodan Liang, Olga Russakovsky, Yao Zhao, Yunchao Wei*
- MILES: Visual BERT Pre-training with Injected Language Semantics for Video-Text Retrieval, *Yuying Ge, Yixiao Ge, Xihui Liu, Jinpeng Wang, Jianping Wu, Ying Shan, Xiaohu Qie, Ping Luo*
- GEB+: A Benchmark for Generic Event Boundary Captioning, Grounding and Retrieval, *Yuxuan Wang, Difei Gao, Licheng Yu, Weixian Lei, Matt Feiszli, Mike Zheng Shou*
- A Simple and Robust Correlation Filtering Method for Text-Based Person Search, *Wei Suo, Mengyang Sun, Kai Niu, Yiqi Gao, Peng Wang, Yanning Zhang, Qi Wu*

- Trace Controlled Text to Image Generation, *Kun Yan, Lei Ji, Chenfei Wu, Jianmin Bao, Ming Zhou, Nan Duan, Shuai Ma*
 - Rethinking Data Augmentation for Robust Visual Question Answering, *Long Chen, Yuhang Zheng, Jun Xiao*
 - Explicit Image Caption Editing, *Zhen Wang, Long Chen, Wenbo Ma, Guangxing Han, Yulei Niu, Jian Shao, Jun Xiao*
 - Can Shuffling Video Benefit Temporal Bias Problem: A Novel Training Framework for Temporal Grounding, *Jiachang Hao, Haifeng Sun, Pengfei Ren, Jingyu Wang, Qi Qi, Jianxin Liao*
 - GRIT: Faster and Better Image Captioning Transformer Using Dual Visual Features, *Van-Quang Nguyen, Masanori Suganuma, Takayuki Okatani*
 - Spatial and Visual Perspective-Taking via View Rotation and Relation Reasoning for Embodied Reference Understanding, *Cheng Shi, Sibe Yang*
 - Contrastive Vision-Language Pre-training with Limited Resources, *Quan Cui, Boyan Zhou, Yu Guo, Weidong Yin, Hao Wu, Osamu Yoshie, Yubo Chen*
 - Learning Linguistic Association towards Efficient Text-Video Retrieval, *Sheng Fang, Shuhui Wang, Junbao Zhuo, Xinzhe Han, Qingming Huang*
 - Switch-BERT: Learning to Model Multimodal Interactions by Switching Attention and Input, *Qingpei Guo, Kaisheng Yao, Wei Chu*
 - Word-Level Fine-Grained Story Visualization, *Bowen Li*
 - Unifying Event Detection and Captioning as Sequence Generation via Pre-training, *Qi Zhang, Yuqing Song, Qin Jin*
 - Fine-Grained Visual Entailment, *Christopher Thomas, Yipeng Zhang, Shih-Fu Chang*
 - New Datasets and Models for Contextual Reasoning in Visual Dialog, *Yifeng Zhang, Ming Jiang, Qi Zhao*
 - AssistQ: Affordance-Centric Question-Driven Task Completion for Egocentric Assistant, *Benita Wong, Joya Chen, You Wu, Stan Weixian Lei, Dongxing Mao, Difei Gao, Mike Zheng Shou*
 - Scaling Open-Vocabulary Image Segmentation with Image-Level Labels, *Golnaz Ghiasi, Xiuye Gu, Yin Cui, Tsung-Yi Lin*
 - TISE: Bag of Metrics for Text-to-Image Synthesis Evaluation, *Tan M. Dinh, Rang Nguyen, Binh-Son Hua*
 - SemAug: Semantically Meaningful Image Augmentations for Object Detection through Language Grounding, *Morgan Heisler, Amin Banitalebi-Dehkordi, Yong Zhang*
 - Referring Object Manipulation of Natural Images with Conditional Classifier-Free Guidance, *Myungsub Choi*
 - FedVLN: Privacy-Preserving Federated Vision-and-Language Navigation, *Kaiwen Zhou, Xin Eric Wang*
 - CODER: Coupled Diversity-Sensitive Momentum Contrastive Learning for Image-Text Retrieval, *Haoran Wang, Dongliang He, Wenhao Wu, Boyang Xia, Min Yang, Fu Li, Yunlong Yu, Zhong Ji, Errui Ding, Jingdong Wang*
 - Language-Driven Artistic Style Transfer, *Tsu-Jui Fu, Xin Eric Wang, William Yang Wang*
 - Single-Stream Multi-level Alignment for Vision-Language Pretraining, *Zaid Khan, Vijay Kumar B G, Xiang Yu, Samuel Schuster, Manmohan Chandraker, Yun Fu*
 - Most and Least Retrievable Images in Visual-Language Query Systems, *Liuwan Zhu, Rui Ning, Jiang Li, Chunsheng Xin, Hongyi Wu*
 - Sports Video Analysis on Large-Scale Data, *Dekun Wu, He Zhao, Xingce Bao, Richard P. Wildes*
 - StoryDALL-E: Adapting Pretrained Text-to-Image Transformers for Story Continuation, *Adyasha Maharana, Darryl Hannan, Mohit Bansal*
- Vision + Other Modalities**
- End-to-End Active Speaker Detection, *Juan León Alcázar, Moritz Cordes, Chen Zhao, Bernard Ghanem*
 - Emotion Recognition for Multiple Context Awareness, *Dingkang Yang, Shuai Huang, Shunli Wang, Yang Liu, Peng Zhai, Liuzhen Su, Mingcheng Li, Lihua Zhang*
 - Adaptive Fine-Grained Sketch-Based Image Retrieval, *Ayan Kumar Bhunia, Aneeshan Sain, Parth Hiren Shah, Animesh Gupta, Pinaki Nath Chowdhury, Tao Xiang, Yi-Zhe Song*
 - Uncertainty-Aware Multi-modal Learning via Cross-Modal Random Network Prediction, *Hu Wang, Jianpeng Zhang, Yuanhong Chen, Congbo Ma, Jodie Avery, Louise Hull, Gustavo Carneiro*
 - Localizing Visual Sounds the Easy Way, *Shentong Mo, Pedro Morgado*
 - Remote Respiration Monitoring of Moving Person Using Radio Signals, *Jae-Ho Choi, Ki-Bong Kang, Kyung-Tae Kim*
 - PACS: A Dataset for Physical Audiovisual Commonsense Reasoning, *Samuel Yu, Peter Wu, Paul Pu Liang, Ruslan Salakhutdinov, Louis-Philippe Morency*
 - Audio—Visual Segmentation, *Jinxing Zhou, Jianyuan Wang, Jiayi Zhang, Weixuan Sun, Jing Zhang, Stan Birchfield, Dan Guo, Lingpeng Kong, Meng Wang, Yiran Zhong*
- Vision Applications and Systems**
- GAMA: Cross-view Video Geo-localization, *Shruti Vyas, Chen Chen, Mubarak Shah*
 - Geometric Representation Learning for Document Image Rectification, *Hao Feng, Wengang Zhou, Jiajun Deng, Yuechen Wang, Houqiang Li*
 - S2-VER: Semi-Supervised Visual Emotion Recognition, *Guoli Jia, Jufeng Yang*
 - Image Coding for Machines with Omnipotent Feature Learning, *Ruoyu Feng, Xin Jin, Zongyu Guo, Runsen Feng, Yixin Gao, Tianyu He, Zhizheng Zhang, Simeng Sun, Zhibo Chen*
 - Fashionformer: A Simple, Effective and Unified Baseline for Human Fashion Segmentation and Recognition, *Shilin Xu, Xiangtai Li, Jingbo Wang, Guangliang Cheng, Yunhai Tong, Dacheng Tao*
 - Relationship Spatialization for Depth Estimation, *Xiaoyu Xu, Jiayan Qiu, Xinchao Wang, Zhou Wang*
 - MegBA: A GPU-Based Distributed Library for Large-Scale Bundle Adjustment, *Jie Ren, Wenteng Liang, Ran Yan, Luo Mai, Shiwen Liu, Xiao Liu*
 - An Efficient Person Clustering Algorithm for Open Checkout-Free Groceries, *Junde Wu, Yu Zhang, Rao Fu, Yuanpei Liu, Jing Gao*
 - Actor-Centered Representations for Action Localization in Streaming Videos, *Sathyanarayanan Aakur, Sudeep Sarkar*
 - Bandwidth-Aware Adaptive Codec for DNN Inference Offloading in IoT, *Xiufeng Xie, Ning Zhou, Wentao Zhu, Ji Liu*
 - Responsive Listening Head Generation: A Benchmark Dataset and Baseline, *Mohan Zhou, Yalong Bai, Wei Zhang, Ting Yao, Tiejun Zhao, Tao Mei*
 - Towards Scale-Aware, Robust, and Generalizable Unsupervised Monocular Depth Estimation by Integrating IMU Motion Dynamics, *Sen Zhang, Jing Zhang, Dacheng Tao*
 - Addressing Heterogeneity in Federated Learning via Distributional Transformation, *Haolin Yuan, Bo Hui, Yuchen Yang, Philippe Burlina, Neil Zhenqiang Gong, Yinzhi Cao*

- Where in the World Is This Image? Transformer-Based Geo-Localization in the Wild, *Shraman Pramanick, Ewa M. Nowara, Joshua Gleason, Carlos D. Castillo, Rama Chellappa*
- Colorization for In Situ Marine Plankton Images, *Guannan Guo, Qi Lin, Tao Chen, Zhenghui Feng, Zheng Wang, Jianping Li*
- Efficient Deep Visual and Inertial Odometry with Adaptive Visual Modality Selection, *Mingyu Yang, Yu Chen, Hun-Seok Kim*
- A Cloud 3D Dataset and Application-Specific Learned Image Compression in Cloud 3D, *Tianyi Liu, Sen He, Vinodh Kumaran Jayakumar, Wei Wang*
- AutoTransition: Learning to Recommend Video Transition Effects, *Yaojie Shen, Libo Zhang, Kai Xu, Xiaojie Jin*

Vision for and Autonomous Vehicles

- Open-World Semantic Segmentation for LIDAR Point Clouds, *Jun Cen, Peng Yun, Shiwei Zhang, Junhao Cai, Di Luan, Mingqian Tang, Ming Liu, Michael Yu Wang*
- InAction: Interpretable Action Decision Making for Autonomous Driving, *Taotao Jing, Haifeng Xia, Renran Tian, Haoran Ding, Xiao Luo, Joshua Domeyer, Rini Sherony, Zhengming Ding*
- CODA: A Real-World Road Corner Case Dataset for Object Detection in Autonomous Driving, *Kaican Li, Kai Chen, Haoyu Wang, Lanqing Hong, Chaoqiang Ye, Jianhua Han, Yukuai Chen, Wei Zhang, Chunjing Xu, Dit-Yan Yeung, Xiaodan Liang, Zhenguo Li, Hang Xu*
- RCLane: Relay Chain Prediction for Lane Detection, *Shenghua Xu, Xinyue Cai, Bin Zhao, Li Zhang, Hang Xu, Yanwei Fu, Xiangyang Xue*
- Physical Attack on Monocular Depth Estimation with Optimal Adversarial Patches, *Zhiyuan Cheng, James Liang, Hongjun Choi, Guanhong Tao, Zhiwen Cao, Dongfang Liu, Xiangyu Zhang*
- ST-P3: End-to-End Vision-Based Autonomous Driving via Spatial-Temporal Feature Learning, *Shengchao Hu, Li Chen, Penghao Wu, Hongyang Li, Junchi Yan, Dacheng Tao*
- PointFix: Learning to Fix Domain Bias for Robust Online Stereo Adaptation, *Kwonyoung Kim, Jungin Park, Jiyoung Lee, Dongbo Min, Kwanghoon Sohn*
- BRNet: Exploring Comprehensive Features for Monocular Depth Estimation, *Wencheng Han, Junbo Yin, Xiaogang Jin, Xiangdong Dai, Jianbing Shen*
- SiamDoGe: Domain Generalizable Semantic Segmentation Using Siamese Network, *Zhenyao Wu, Xinyi Wu, Xiaoping Zhang, Lili Ju, Song Wang*
- Multimodal Transformer for Automatic 3D Annotation and Object Detection, *Chang Liu, Xiaoyan Qian, Binxiao Huang, Xiaojuan Qi, Edmund Lam, Siew-Chong Tan, Ngai Wong*
- Homogeneous Multi-modal Feature Fusion and Interaction for 3D Object Detection, *Xin Li, Botian Shi, Yuenan Hou, Xingjiao Wu, Tianlong Ma, Yikang Li, Liang He*
- JPerceiver: Joint Perception Network for Depth, Pose and Layout Estimation in Driving Scenes, *Haimei Zhao, Jing Zhang, Sen Zhang, Dacheng Tao*
- Semi-Supervised 3D Object Detection with Proficient Teachers, *Junbo Yin, Jin Fang, Dingfu Zhou, Liangjun Zhang, Cheng-Zhong Xu, Jianbing Shen, Wenguan Wang*
- Point Cloud Compression with Sibling Context and Surface Priors, *Zhili Chen, Zian Qian, Sukai Wang, Qifeng Chen*
- Lane Detection Transformer Based on Multi-Frame Horizontal and Vertical Attention and Visual Transformer Module, *Han Zhang, Yunchao Gu, Xinliang Wang, Junjun Pan, Minghui Wang*

- ProposalContrast: Unsupervised Pre-training for LiDAR-Based 3D Object Detection, *Junbo Yin, Dingfu Zhou, Liangjun Zhang, Jin Fang, Cheng-Zhong Xu, Jianbing Shen, Wenguan Wang*
- PreTraM: Self-Supervised Pre-training via Connecting Trajectory and Map, *Chenfeng Xu, Tian Li, Chen Tang, Lingfeng Sun, Kurt Keutzer, Masayoshi Tomizuka, Alireza Fathi, Wei Zhan*
- V2X-ViT: Vehicle-to-Everything Cooperative Perception with Vision Transformer, *Runsheng Xu, Hao Xiang, Zhengzhong Tu, Xin Xia, Ming-Hsuan Yang, Jiaqi Ma*
- DevNet: Self-Supervised Monocular Depth Learning via Density Volume Construction, *Kaichen Zhou, Lanqing Hong, Changhao Chen, Hang Xu, Chaoqiang Ye, Qingyong Hu, Zhenguo Li*
- Radatron: Accurate Detection Using Multi-Resolution Cascaded MIMO Radar, *Sohrab Madani, Jayden Guan, Waleed Ahmed, Saurabh Gupta, Haitham Hassanieh*
- LiDAR Distillation: Bridging the Beam-Induced Domain Gap for 3D Object Detection, *Yi Wei, Zibu Wei, Yongming Rao, Jiaxin Li, Jie Zhou, Jiwen Lu*
- Efficient Point Cloud Segmentation with Geometry-Aware Sparse Networks, *Maosheng Ye, Rui Wan, Shuangjie Xu, Tongyi Cao, Qifeng Chen*
- Rethinking Closed-Loop Training for Autonomous Driving, *Chris Zhang, Runsheng Guo, Wenyuan Zeng, Yuwen Xiong, Binbin Dai, Rui Hu, Mengye Ren, Raquel Urtasun*

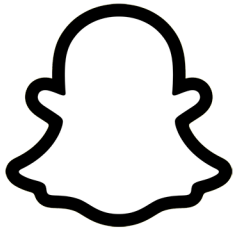
Vision for Robotics and Embodied Vision

- Generative Meta-Adversarial Network for Unseen Object Navigation, *Sixian Zhang, Weijie Li, Xinhang Song, Yubing Bai, Shuqiang Jiang*
- Object Manipulation via Visual Target Localization, *Kiana Ehsani, Ali Farhadi, Aniruddha Kembhavi, Roozbeh Mottaghi*
- Domain Randomization-Enhanced Depth Simulation and Restoration for Perceiving and Grasping Specular and Transparent Objects, *Qiyu Dai, Jiayao Zhang, Qiwei Li, Tianhao Wu, Hao Dong, Ziyuan Liu, Ping Tan, He Wang*
- Resolving Copycat Problems in Visual Imitation Learning via Residual Action Prediction, *Chia-Chi Chuang, Donglin Yang, Chuan Wen, Yang Gao*
- TransGrasp: Grasp Pose Estimation of a Category of Objects by Transferring Grasps from Only One Labeled Instance, *Hongtao Wen, Jianhang Yan, Wanli Peng, Yi Sun*
- StARformer: Transformer with State-Action-Reward Representations for Visual Reinforcement Learning, *Jinghuan Shang, Kumara Kahatapitiya, Xiang Li, Michael S. Ryoo*
- Learning Efficient Multi-Agent Cooperative Visual Exploration, *Chao Yu, Xinyi Yang, Jiaxuan Gao, Huazhong Yang, Yu Wang, Yi Wu*
- Sim-to-Real 6D Object Pose Estimation via Iterative Self-Training for Robotic Bin Picking, *Kai Chen, Rui Cao, Stephen James, Yichuan Li, Yun-Hui Liu, Pieter Abbeel, Qi Dou*
- Self-Supervised Interactive Object Segmentation through a Singulation-and-Grasping Approach, *Houjian Yu, Changhyun Choi*
- Learning from Unlabeled 3D Environments for Vision-and-Language Navigation, *Shizhe Chen, Pierre-Louis Guhur, Makarand Tapaswi, Cordelia Schmid, Ivan Laptev*

Visual Reasoning and Logical Representation

- Learning Algebraic Representation for Systematic Generalization in Abstract Reasoning, *Chi Zhang, Sirui Xie, Baoxiong Jia, Ying Nian Wu, Song-Chun Zhu, Yixin Zhu*

Gold Donors



Silver Donors



Exhibitors



Startup Exhibitors



Media Sponsors





BOSCH



HUAWEI

