

IEEE/CVF Conference on Computer Vision and Pattern Recognition

Program Guide
Main Conference

JUNE 18-22, 2023

CVPR



VANCOUVER, CANADA

PLATINUM SPONSORS

amazon | science



cruise

Google



Qualcomm



TOYOTA
RESEARCH INSTITUTE

GOLD SPONSORS

Alibaba Cloud



datagen

FURIOSA

LATITUDE



NOVARC
TECHNOLOGIES

speechocean

Synthesis AI

TELUS[®] International

TESLA

TikTok

Weights & Biases

ZOOX

SILVER SPONSORS

Dataminr[®]



HYUNDAI

kitware

LG AI Research

Lightning^{AI}

manot

meitu | M² Lab
Meitu Imaging & Vision Lab

PROPHESSEE
METAVISION FOR MACHINES

RIVIAN

scale



CHEN^{TIANQIAO & CHRISSY}
INSTITUTE

Visual Layer

WAYMO

Message from the General and Program Chairs

Welcome to the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition in Vancouver, Canada. As in previous years, CVPR is the premier and flagship annual meeting of IEEE/CVF and PAMI-TC, where researchers in our community present their latest advances in computer vision, pattern recognition, machine learning, robotics, and artificial intelligence, both in theory and practice. Our program includes invited keynote talks, award paper presentations, poster presentations, tutorials, workshops, demos, exhibitions, and an amiable social setting, all aimed at providing attendees with an exciting and enriching experience.

This year marks the first time in a while that many pandemic restrictions have been lifted, allowing us to come together in person again to celebrate the latest advances in our field. For those unable to join us physically, we are pleased to offer a virtual component that will provide access to conference papers, posters, videos, and talks. We hope this virtual option will allow everyone to engage with the exciting research being presented.

CVPR 2023 received 9155 submissions, a 12% increase from the 8161 submissions to CVPR 2022. The review process was managed by 400+ area chairs and, new to the process this year, 30 senior area chairs. The senior area chairs helped in a number of respects, most importantly by adjudicating difficult cases, covering emergencies, selecting highlight papers, and selecting the award candidates. During the review phase, each paper received at least 3 reviews from the pool of 6625 reviewers. As in prior years, after receiving these initial reviews, the process continued with an author rebuttal phase, discussion among reviewers and ACs, finalizing of reviews, and ACs working in triplets to make final accept/reject decisions for each paper. At the end of this process, 2359 papers were accepted (25.8% acceptance rate). In keeping with the CVPR tradition, the PCs did not pre-set any acceptance cap. The resulting acceptance rate reflects the community consensus, and is well aligned with past CVPRs.

Of the 2359 accepted papers, 235 (10%) were selected as highlights. In addition, 12 (0.5%) papers have been shortlisted as best paper award candidates. The final best papers and honorable mentions are selected from these 12 papers by an independent award committee appointed by the program chairs, which is composed of experienced researchers from our community. The award committee is led by an award committee chair appointed by the program chairs, who moderates the selection process.

This year, CVPR will be single-track to allow everyone to attend everything. The focus will be on a few plenary talks, keynotes and panels, and plenty of time for poster sessions, networking, and socializing. Every paper will be presented at a poster session. All paper award candidates will have an additional plenary oral presentation. Every attendee will have access to a personalized digital program to easily navigate the ~400 posters in each poster session. The virtual platform will host papers, posters, videos, and a chat room for every paper. All plenary events will be streamed online for all attendees that cannot attend in person.

We would like to thank everyone involved in making CVPR 2023 a success. This includes the organizing committee, area chairs, senior area chairs, reviewers, authors, demo session participants, donors, exhibitors, and everyone else without whom this meeting would not be possible. We also thank Nicole Finn and her C to C Events team for organizing the conference logistics, Lee Campbell and the Event Hosts team for their work on the website and virtual platform, and Mike Weil and Hall Erickson for handling sponsorships and the exhibition. Last but not least, we thank all of you for attending CVPR 2023 and making it one of the top venues for computer vision research in the world. We hope that you also have some time to explore gorgeous Vancouver during the conference. Enjoy CVPR 2023. We look forward to meeting you in person!

Program Chairs: Andreas Geiger, Ross Girshick, Judy Hoffman, Vladlen Koltun, and Svetlana Lazebnik

General Chairs: Michael S. Brown, Fei-Fei Li, Greg Mori, and Yoichi Sato

CVPR 2023 QR Codes

CVPR 2023 Virtual Platform



Access schedules, papers, workshops, tutorials, etc.

CVPR 2023 Slido Site



Attendee engagement (Q&A, polls, etc.) for plenary sessions: keynotes, panels, orals.

CVPR 2023 Organizing Committee

General Chairs:	Michael S. Brown Fei-Fei Li Greg Mori Yoichi Sato	Diversity, Equity, & Inclusion Chairs:	Thibaut Durand Fatma Güney Kate Saenko
Program Chairs:	Andreas Geiger Ross Girshick Judy Hoffman Vladlen Koltun Svetlana Lazebnik	Publicity Chairs:	Kosta Derpanis Boqing Gong Abby Stylianou
Workshops Chairs:	Olga Russakovsky Yu Wu Serena Yeung	Social Chair:	Yale Song
Tutorials Chairs:	Siyu Tang Jianxin Wu	Social Activities Chairs:	Angel Chang Giovanni Farinella
Demonstrations Chairs:	Jon Barron Gim Hee Lee	Local Arrangements Chairs:	Kwang Moo Yi Leon Sigal
Doctoral Consortium Chairs:	'YZ' Yezhou Yang Catherine Qi Zhao	Virtual Platform Chair:	Andreas Geiger
Program Advisory Board:	Michael J. Black David Forsyth Kristen Grauman Jitendra Malik Cordelia Schmid Richard Szeliski Andrew Zisserman	Accessibility Chair:	Danna Gurari
		Finance Chair:	Bryan Morse
		Technical Chair:	David Hafner
		Webmaster:	Lee Campbell
		Senior PAMI-TC Ombuds:	David Forsyth Linda Shapiro
		CVPR 2023 Ombuds:	Kyoung Mu Lee Xiaodan Liang
		Publications Chair:	Eric Mortensen
		Event Producer:	Nicole Bumpus Finn

CVPR 2023 Senior Area Chairs

Serge Belongie	Derek Hoiem	Stephen Lin	Stefan Roth
Alexander C. Berg	Gang Hua	Zhouchen Lin	Kate Saenko
Dima Damen	C.V. Jawahar	Simon Lucey	Cordelia Schmid
Piotr Dollar	Herve Jegou	Laurens van der Maaten	Tinne Tuytelaars
David Forsyth	Seon Joo Kim	Yasuyuki Matsushita	Ming-Hsuan Yang
Georgia Gkioxari	Kyros Kutulakos	Ko Nishino	Jingyi Yu
Abhinav Gupta	Bastian Leibe	Wanli Ouyang	
Bohyung Han	Hongdong Li	Ian Reid	

CVPR 2023 Area Chairs

Sathyanarayanan N. Aakur	Adrien Bartoli	Angel X Chang	Cheng Deng	Katerina Fragkiadaki
Ehsan Abbasnejad	Rodrigo Benenson	Huiwen Chang	Jia Deng	Alejandro F. Frangi
Ehsan Adeli	Siavash Arjomand Bigdeli	Anpei Chen	Hao Dong	Pascal Fua
Zeynep Akata	Soma Biswas	Dong Chen	Matthijs Douze	Yasutaka Furukawa
Karteek Alahari	Matthew B. Blaschko	Dongdong Chen	Michal Drozdal	Li Fuxin
Alexandre Alahi	Piotr Bojanowski	Hwann-Tzong Chen	Thibaut Durand	Fabio Galasso
Mohsen Ali	Edmond Boyer	Liang-Chieh Chen	Mohamed Elhoseiny	Juergen Gall
Rahaf Aljundi	Marcus A Brubaker	Long Chen	Sergio Escalera	Chuang Gan
Kalle Åström	Xilin CHEN	Ming-Syan Chen	Bin Fan	Zhe Gan
Oisin Mac Aodha	Zhaowei Cai	Qifeng Chen	Deng-Ping Fan	Shenghua Gao
Andre Araujo	Neill D. F. Campbell	Shizhe Chen	Giovanni Maria Farinella	Efstratios Gavves
Pablo Arbelaez	Xiaochun Cao	Xinlei Chen	Alireza Fathi	Peter Vincent Gehler
Anurag Arnab	Gustavo Carneiro	Ming-Ming Cheng	Paolo Favaro	Theo Gevers
Chetan Arora	Joao Carreira	Yu Cheng	Christoph Feichtenhofer	Golnaz Ghiasi
Yuki M Asano	Ayan Chakrabarti	Minsu Cho	Michael Felsberg	Rohit Girdhar
Mathieu Aubry	Ishani Chakraborty	Jonghyun Choi	Rogério Feris	Ioannis Gkioulekas
R. Venkatesh Babu	Shayok Chakraborty	Yung-Yu Chuang	Vittorio Ferrari	Boqing Gong
Song Bai	Tat-Jen Cham	John Collomosse	Cristian Canton Ferrer	Stephen Gould
Lamberto Ballan	Antoni B. Chan	Jifeng Dai	David Fouhey	David Gu
Joao P Barreto	Manmohan Chandraker	Yuchao Dai	Charles Fowlkes	Jinwei Gu

CVPR 2023 Area Chairs (cont.)

Fatma Guney	Philipp Kraehenbuehl	Vinay P Namboodiri	Viktoriia Sharmanska	Xiaoyu Wang
Yulan Guo	Matej Kristan	Ram Nevatia	Evan Shelhamer	Yu-Xiong Wang
Mohit Gupta	Hilde Kuehne	Natalia Neverova	Wei Shen	Gordon Wetzstein
Saurabh Gupta	Zuzana Kukelova	Lam M. Nguyen	Mike Zheng Shou	Lior Wolf
Tanmay Gupta	Suha Kwak	Bingbing Ni	Abhinav Shrivastava	Kwan-Yee K. Wong
Junwei Han	Shang-Hong Lai	Shohei Nobuhara	Leonid Sigal	Michael Wray
Tian Han	Cuiling Lan	David Novotny	Saurabh Singh	Chao-Yuan Wu
Xiaoguang Han	Laura Leal-Taixé	Matthew O'Toole	Vincent Sitzmann	Jianxin Wu
Zhizhong Han	Hsin-Ying Lee	Takayuki Okatani	Noah Snaveley	Qi Wu
Emily Morgan Hand	Yong Jae Lee	Björn Ommer	Cees G. M. Snoek	Tianfu Wu
Tatsuya Harada	Vincent Lepetit	Vicente Ordóñez	Kihyuk Sohn	Yi Wu
Tal Hassner	Bo Li	Andrew Owens	Yang Song	Yihong Wu
Di He	Hongsheng Li	Jinshan Pan	Yibing Song	Yu Wu
Ran He	Hongyang Li	Dim Papadopoulos	Richard Souvenir	Zhirong Wu
Xuming He	Nianyi Li	Jaesik Park	Pratul P. Srinivasan	Zuxuan Wu
Aaron Hertzmann	Xueting Li	Vishal M. Patel	Abby Stylianou	Chaowei Xiao
Otmar Hilliges	Yijun Li	Deepak Pathak	Bing Su	Jianwen Xie
Seunghoon Hong	Yin Li	Ioannis Patras	Yu-Chuan Su	Weidi Xie
Anthony Hoogs	Yu Li	Marcello Pelillo	Waqas Sultani	Cihang Xie
Yedid Hoshen	Zhouhui Lian	Hamed Pirsiavash	Chen Sun	Chang Xu
Mahdi S. Hosseini	Hong-Yuan Mark Liao	Robert Pless	Deqing Sun	Dan Xu
Jia-Bin Huang	Shengcai Liao	Bryan A. Plummer	Min Sun	Huijuan Xu
Qixing Huang	Yiyi Liao	Gerard Pons-Moll	Andrea Tagliasacchi	Xiangyu Xu
Sharon X Huang	Ser-Nam Lim	Jordi Pont-Tuset	Yu-Wing Tai	Tianfan Xue
Wei-Chih Hung	Liang Lin	Brian L. Price	Mingkui Tan	Toshihiko Yamasaki
Eddy Ilg	Tsung-Yi Lin	Xiaojuan Qi	Robby T. Tan	Jimei Yang
Phillip Isola	Yen-Yu Lin	Yu Qiao	Camillo Jose Taylor	Ruigang Yang
Nathan Jacobs	David B. Lindell	Ravi Ramamoorthi	Yingli Tian	Wei Yang
Unnat Jain	Ce Liu	Deva Ramanan	Joseph Tighe	Qixiang Ye
Stephen James	Risheng Liu	Rene Ranftl	Radu Timofte	Raymond A. Yeh
Varun Jampani	Si Liu	Tammy Riklin Raviv	Federico Tombari	Sai-Kit Yeung
Qiang Ji	Ziwei Liu	Wenqi Ren	Tatiana Tommasi	Serena Yeung
Huaizu Jiang	miaomiao Liu	Hamid Rezatofighi	Xin Tong	Kwang Moo Yi
Lu Jiang	Stephen Lombardi	Helge Rhodin	Fernando De la Torre	Li Yi
Justin Johnson	Chengjiang Long	Stephan R. Richter	Lorenzo Torresani	Shaodi You
Hanbyul Joo	Marco Lorenzi	Tobias Ritschel	Alexander T Toshev	Fisher Yu
Jungseock Joo	Chen Change Loy	Emanuele Rodolà	Anh Tuan Tran	Ning Yu
Neel Joshi	Cewu Lu	Amit Roy-Chowdhury	Yi-Hsuan Tsai	Stella X. Yu
Da-Cheng Juan	Feng Lu	Christian Rupprecht	Hung-Yu Tseng	Yizhou Yu
Fredrik Kahl	Chong Luo	Olga Russakovsky	Zhuowen Tu	Lu Yuan
Yannis Kalantidis	Ping Luo	Bryan Russell	Shubham Tulsiani	Amir Zamir
Evangelos Kalogerakis	Khoa Luu	Ken Sakurada	Sergey Tulyakov	Guofeng Zhang
Angjoo Kanazawa	Dhruv Mahajan	Mathieu Salzmann	Tony Tung	Li Zhang
Asako Kanezaki	Michael Maire	Dimitris Samaras	Georgios Tzimiropoulos	Muhan Zhang
Sing Bing Kang	Subhransu Maji	Aswin C. Sankaranarayanan	Jasper Uijlings	Ning Zhang
Dimosthenis Karatzas	Yasushi Makihara	Sudeep Sarkar	Norimichi Ukita	Richard Zhang
Hiroshi Kawasaki	Arun Mallya	Imari Sato	Maria Vakalopoulou	Tianzhu Zhang
Aniruddha Kembhavi	Jiri Matas	Torsten Sattler	Jack Valmadre	Xiangyu Zhang
Margret Keuper	Stefano Mattoccia	Manolis Savva	Gül Varol	Yulun Zhang
Fahad Khan	Shalini De Mello	Daniel Scharstein	Nuno Vasconcelos	Zhaoxiang Zhang
Salman Khan	Pascal Mettes	Walter Scheirer	Andrea Vedaldi	Hengshuang Zhao
Dahun Kim	Peyman Milanfar	Bernt Schiele	Chaohui Wang	Qi Zhao
Gunhee Kim	Ben Mildenhall	Konrad Schindler	Chen Wang	Liang Zheng
Min H. Kim	Ishan Misra	Julia A Schnabel	Chunyu Wang	Wei-Shi Zheng
Seungryong Kim	Niloy Mitra	Alex Schwing	Jian Wang	Yinqiang Zheng
Zolt Kira	Philippos Mordohai	Ozan Sener	Jingdong Wang	Jiahuan Zhou
Alexander Kirillov	Francesc Moreno-Noguer	Paul Hongsuck Seo	Jingya Wang	S Kevin Zhou
Laurent Kneip	Pedro Morgado	Thomas Serre	Limin Wang	Yixin Zhu
Nikos Komodakis	Roozbeh Mottaghi	Fei Sha	Liwei Wang	Chuhang Zou
Piotr Koniusz	Vittorio Murino	Greg Shakhnarovich	Oliver Wang	Silvia Zuffi
Adriana Kovashka	Arsha Nagrani	Shiguang Shan	Shenlong Wang	

CVPR 2023 Outstanding Reviewers

We are pleased to recognize the following researchers as “CVPR 2023 Outstanding Reviewers”. These reviewers contributed reviews noted as excellent by area chairs and are being recognized for their outstanding community service.

Abhinav Shukla	Fangzhou Mu	Kewei Wang	Peipei Li	Williem Williem
Alessio Sampieri	Federico Perazzi	Kha Gia Quach	Pirazh Khorramshahi	Wouter Van Gansbeke
Aliaksandr Siarohin	Francesco Cappio Borlino	Koki Nagano	Piyush Nitin Bagad	Xiaohan Ding
Ameesh Makadia	Francesco Croce	Krystian Mikolajczyk	Prateek Prasanna	Xiaohan Wang
Amit Raj	Gaku Nakano	Kunal Swami	Prithvijit Chattopadhyay	Xiaolong Li
Anand Mishra	Gilles Puy	Kushal Kafle	Priyam Dey	Xiaoqing Guo
Andrea Burns	Guillermo Gallego	Li Chen	Qi Bi	Xiaoyang Wu
Andrei Bursuc	Gukyeong Kwon	Li Niu	Qiongjie Cui	Xiaoyi Dong
Andrew Zisserman	Gyeongsik Moon	Long Sun	Relja Arandjelovic	Xin Lai
Apoorv Khandelwal	Hannah Bull	Luca Cosmo	Renato Martins	Xintong Han
Arif Mahmood	Haoran Bai	Luca Weihs	Richard J. Chen	Xu Chen
Axel Barroso-Laguna	Haotong Qin	Lucas Beyer	Rohan Chabra	Xu Ma
Baoyuan Wu	Hazel Doughty	M. Saquib Sarfraz	Romain Brégier	Xuan Bac Nguyen
Bernhard Zeisl	Henghui Ding	M. Usman Rafique	Ronghang Hu	Xucong Zhang
Bin Xia	Hongbo Fu	Mandi Luo	Rémi Pautrat	Yan Huang
Binh-Son Hua	Hossein Rahmani	Marcel Buehler	Samuele Salti	Yan Zhao
Bruno Korbar	Huan Liu	Marco Cristani	Samyak Jain	Yang He
Carlo Masone	Huan Wang	Marco Fumero	Sara Beery	Yangyan Li
Chaoyou Fu	Huanrui Yang	Marcus Rohrbach	Scott Wisdom	Yannick Strümpfer
Chen Song	Hung-Jin Lin	Maria A Zuluaga	Seonwook Park	Yasamin Jafarian
Chi Xu	Idil Esen Zulfikar	Mark Sheinin	Sergey Prokudin	Yi Li
Chi Zhang	Ioana Croitoru	Martin Weinmann	Shengcao Cao	Yifan Liu
Chonghao Sima	Iro Laina	Mathis Petrovich	Shengju Qian	Yihua Zhang
Chongjian GE	Itai Lang	Matt Deitke	Shuai Yang	Yuanhao Cai
Chongyi Li	Ivan Skorokhodov	Matteo Poggi	Shuyang Gu	Yumin Suh
Chris Rockwell	Jaime Spencer	Matteo Ruggero Ronchi	Simon Niklaus	Yunzhong Hou
Christian Richardt	Jamie Watson	Matthew Chan	Simone Melzi	Yusuke Matsui
Christian Unger	Janne Heikkilä	Melinos Averkiou	Simone Schaub-Meyer	Yutong Xie
Christopher Thomas	Jason Saragih	Mengtian Li	Soonmin Hwang	Zhan Tong
Chuanxia Zheng	Jens Behley	Michael J. Black	Spyros Gidaris	Zheng Xu
Constantin Pape	Jerome Revaud	Michael Niemeyer	Stamatios Georgoulis	Zhenzhi Wang
Dailan He	Jiahao Xie	Michael Waechter	Stergios Christodoulidis	Ziad Al-Halah
Daksh Thapar	Jiangxin Dong	Mina Ghadimi Atigh	Suhas Lohit	Zikui Cai
Daniel Duckworth	Jianmin Bao	Minesh Mathew	Sukrut Rao	Zixiang Zhao
David Picard	Jiawei Zhang	Mingchen Zhuge	Supasorn Suwajanakorn	Ziyi Wang
David T Hoffmann	Jiayi Ma	Mingdeng Cao	Suvaansh Bhambri	Zuoyue Li
Despoina Paschalidou	Jie Cao	Muhammad Haris Khan	Taras Kucherenko	
Di Huang	Jindong Gu	Mutian Xu	Tatsunori Tanai	
Dimitrios Tzionas	Jinlong Yang	Nan Xue	Thomas Eboli	
Dmitry P. Vetrov	Josep Lladós	Naoto Inoue	Thomas Fel	
Dmytro Mishkin	Joseph Boyd	Naveen Venkat	Timo Bolkart	
Doyup Lee	Juan Camilo Perez	Neal Wadhwa	Tomas Jakab	
Edoardo Remelli	Julian Ost	Neel Dey	Tomas Vojir	
Eduard Trulls	Jyh-Jing Hwang	Nico Lang	Tony Ng	
Eldar Insafutdinov	Kacper Kania	Oleg Voynov	True Price	
Evangelos Kazakos	Karan Desai	Orchid Majumder	Valentin Gabeur	
Evgenia Rusak	Karsten Roth	Pablo Garrido	Vivien Sainte Fare Garnot	
Fabio Tosi	Kashyap Chitta	Paola Cascante-Bonilla	Weinan Song	
Fangzhou Hong	Kevin J Liang	Paul Albert	Weiwei Sun	

Tuesday, June 20

0700–1700 Registration (West Ballroom Foyer)

0700–0900 Breakfast (West Ballrooms A–D)

0800–0830 Poster Setup (West Exhibit Hall)

0830–0900 Welcome & Opening Remarks (East Exhibit Halls A–B)

0900–1000 Plenary-Tue-AM (East Exhibit Halls A–B)

Keynote: Rodney Brooks (*MIT*)

Title: Revisiting Old Ideas With Modern Hardware

Abstract: Many computer vision ideas have been revisited again and again and again, including current modern computer vision based on neural computation. This round has led to incredible developments in computational hardware. Might such powerful computation breathe life into older neglected ideas?

Chair: Vladlen Koltun (*Apple*)

1000–1800 Exhibits (West Exhibit Hall)

- See Exhibits map for list of exhibitors.

1000–1230 Demos (West Exhibit Hall Demo Area)

- DATID-3D: Diversity-Preserved Domain Adaptation Using Text-to-Image Diffusion for 3D Generative Model, *Gwanghyun Kim, Se Young Chun*
- Deep Lucky Imaging for Astrophotography, *Suraj Singh, Alex Gerko, Dmitry Dylov*
- Fitness Coaching Using LLMs Grounded in Real-Time Vision, *Sunny Panchal, Apratim Bhattacharyya, Mingu Lee, Pulkit Madan, Reza Pourreza, Ingo Bax, Guillaume Berger, Antoine Mercier, Mark Todorovich, Roland Memisevic*
- Contemplation on the Modern Art Through Comparing Them to AI Generated Art, *Sooyeon 'Sue' Park, Subin Kim*
- Streamlining Quality Control: A Guide to Automated Defect Detection With Anomalib, *Paula Ramos, Samet Akcay, Zhuo Wu*
- A Probabilistic Attention Model With Occlusion-Aware Texture Regression for 3D Hand Reconstruction From a Single RGB Image, *Zheheng Jiang, Bryan Williams, Hossein Rahmani*
- Interactive X-Decoder for Understanding and Generating Pixel, Image, and Language, *Xueyan Zou, Jianwei Yang, Yong Jae Lee, Jianfeng Gao*
- On-Device Stable Diffusion on a Snapdragon-Powered Android Smartphone, *Risheek Garrepalli, Kanghai Jang, Carl Blacklock, Markus Nagel, Minseop Park, Jaeseong You, Raghav Srinivasan, Tushar Singhal, Hariharan Sukumaran, Harsh Shah, Liang Zhang, Kim-Chyan Gan, Frank Mayer, Abhijeet Khobare, Mike Tremaine, Jihad Masri, David Callahan, Pooja Sinha, Chirag Patel, Joseph Soriaga, Fatih Porikli*
- OneFormer: One Transformer to Rule Universal Image Segmentation, *Jitesh Jain, Jiachen Li, Mang Tik Chiu, Ali Hassani, Nikita Orlov, Humphrey Shi*

- Applications of NeRF in Unreal Engine, *Fernando Rivas-Manzanique*
- Back to the Source: Diffusion-Driven Adaptation to Corruption, *Dequan Wang, Evan Shelhamer*
- Content Based Search for Deep Generative Models, *Daohan Lu, Sheng-Yu Wang, Nupur Kumari, Rohan Agarwal, Mia Tang, David Bau, Jun-Yan Zhu*
- DINOv2 – Universal Visual Features From Self-Supervised Pre-Training, *Marc Szafraniec, Patrick Labatut, Piotr Bojanowski*
- NUWA, *Chenfei Wu, Nan Duan, Yu Liu, Yang Ou, Yan Xia*
- 3D Cinemagraphy From a Single Image, *Xingyi Li, Zhiguo Cao, Huiqiang Sun, Jianming Zhang, Ke Xian, Guosheng Lin*
- 3D Hand Mesh Estimation at 300 FPS on a Mobile Device, *Jeff X. Zhu, Anurag Ranjan, James Gabriel, Pavan Kumar Anasosalu Vasu, Oncel Tuzel*
- 3D Video Loops From Asynchronous Input, *Li Ma*
- SPOTER Sign Language Recognition in the Browser, *Matyáš Boháček*
- Augmented Reality Instruction for LEGO Arc de Triomphe Construction, *Wei Yan*

1000–1100 Morning Break (West Exhibit Hall)

1030–1230 Poster-Tue-AM (West Exhibit Hall)

✧ - Highlight paper (check it out)

🏆 - Award candidate paper (see award sessions)

1. Megahertz Light Steering Without Moving Parts, *Adithya Pediredla, Srinivasa G. Narasimhan, Maysamreza Chamanzar, Ioannis Gkioulekas*
2. Robust Dynamic Radiance Fields, *Yu-Lun Liu, Chen Gao, Andréas Meuleman, Hung-Yu Tseng, Ayush Saraf, Changil Kim, Yung-Yu Chuang, Johannes Kopf, Jia-Bin Huang*
3. DBARF: Deep Bundle-Adjusting Generalizable Neural Radiance Fields, *Yu Chen, Gim Hee Lee*
4. VDN-NeRF: Resolving Shape-Radiance Ambiguity via View-Dependence Normalization, *Bingfan Zhu, Yanchao Yang, Xulong Wang, Youyi Zheng, Leonidas Guibas*
5. AligNeRF: High-Fidelity Neural Radiance Fields via Alignment-Aware Training, *Yifan Jiang, Peter Hedman, Ben Mildenhall, DeJia Xu, Jonathan T. Barron, Zhangyang Wang, Tianfan Xue*
6. SeaThru-NeRF: Neural Radiance Fields in Scattering Media, *Deborah Levy, Amit Peleg, Naama Pearl, Dan Rosenbaum, Derya Akkaynak, Simon Korman, Tali Treibitz*
7. Exact-NeRF: An Exploration of a Precise Volumetric Parameterization for Neural Radiance Fields, *Brian K. S. Isaac-Medina, Chris G. Willcocks, Toby P. Breckon*
8. Neural Residual Radiance Fields for Streamably Free-Viewpoint Videos, *Liao Wang, Qiang Hu, Qihan He, Ziyu Wang, Jingyi Yu, Tinne Tuytelaars, Lan Xu, Minye Wu*
9. PlenVDB: Memory Efficient VDB-Based Radiance Fields for Fast Training and Rendering, *Han Yan, Celong Liu, Chao Ma, Xing Mei*
10. Local Implicit Ray Function for Generalizable Radiance Field Representation, *Xin Huang, Qi Zhang, Ying Feng, Xiaoyu Li, Xuan Wang, Qing Wang*
11. SurfelNeRF: Neural Surfel Radiance Fields for Online Photorealistic Reconstruction of Indoor Scenes, *Yiming Gao, Yan-Pei Cao, Ying Shan*

12. Frequency-Modulated Point Cloud Rendering With Easy Editing, ✧ Yi Zhang, Xiaoyang Huang, Bingbing Ni, Teng Li, Wenjun Zhang
13. HexPlane: A Fast Representation for Dynamic Scenes, Ang Cao, Justin Johnson
14. Differentiable Shadow Mapping for Efficient Inverse Graphics, Markus Worchel, Marc Alexa
15. Hybrid Neural Rendering for Large-Scale Scenes With Motion Blur, Peng Dai, Yinda Zhang, Xin Yu, Xiaoyang Lyu, Xiaojuan Qi
16. TensoIR: Tensorial Inverse Rendering, Haian Jin, Isabella Liu, Peijia Xu, Xiaoshuai Zhang, Songfang Han, Sai Bi, Xiaowei Zhou, Zexiang Xu, Hao Su
17. ShadowNeuS: Neural SDF Reconstruction by Shadow Ray Supervision, Jingwang Ling, Zhibo Wang, Feng Xu
18. Realistic Saliency Guided Image Enhancement, S. Mahdi H. Miangoleh, Zoya Bylinskii, Eric Kee, Eli Shechtman, Yağiz Aksoy
19. LightPainter: Interactive Portrait Relighting With Freehand Scribble, Yiqun Mei, He Zhang, Xuaner Zhang, Jianming Zhang, Zhixin Shu, Yilin Wang, Zijun Wei, Shi Yan, HyunJoon Jung, Vishal M. Patel
20. A Unified Spatial-Angular Structured Light for Single-View Acquisition of Shape and Reflectance, Xianmin Xu, Yuxin Lin, Haoyang Zhou, Chong Zeng, Yaxin Yu, Kun Zhou, Hongzhi Wu
21. Learning Visibility Field for Detailed 3D Human Reconstruction and Relighting, Ruichen Zheng, Peng Li, Haoqian Wang, Tao Yu
22. Unsupervised Contour Tracking of Live Cells by Mechanical and Cycle Consistency Losses, Junbong Jang, Kwonmoo Lee, Tae-Kyun Kim
23. NeUDF: Leaning Neural Unsigned Distance Fields With Volume Rendering, Yu-Tao Liu, Li Wang, Jie Yang, Weikai Chen, Xiaoxu Meng, Bo Yang, Lin Gao
24. NeAT: Learning Neural Implicit Surfaces With Arbitrary Topologies From Multi-View Images, Xiaoxu Meng, Weikai Chen, Bo Yang
25. ALTO: Alternating Latent Topologies for Implicit 3D Reconstruction, Zhen Wang, Shijie Zhou, Jeong Joon Park, Despoina Paschalidou, Suya You, Gordon Wetzstein, Leonidas Guibas, Achuta Kadambi
26. Controllable Mesh Generation Through Sparse Latent Point Diffusion Models, Zhaoyang Lyu, Jinyi Wang, Yuwei An, Ya Zhang, Dahua Lin, Bo Dai
27. Power Bundle Adjustment for Large-Scale 3D Reconstruction, Simon Weber, Nikolaus Demmel, Tin Chon Chan, Daniel Cremers
28. Neural Pixel Composition for 3D-4D View Synthesis From Multi-Views, Aayush Bansal, Michael Zollhöfer
29. Magic3D: High-Resolution Text-to-3D Content Creation, ✧ Hsuan Lin, Jun Gao, Luming Tang, Towaki Takikawa, Xiaohui Zeng, Xun Huang, Karsten Kreis, Sanja Fidler, Ming-Yu Liu, Tsung-Yi Lin
30. 3D Video Loops From Asynchronous Input, Li Ma, Xiaoyu Li, Jing Liao, Pedro V. Sander
31. High-Fidelity 3D GAN Inversion by Pseudo-Multi-View Optimization, Jiaxin Xie, Hao Ouyang, Jingtian Piao, Chenyang Lei, Qifeng Chen
32. Lift3D: Synthesize 3D Training Data by Lifting 2D GAN to 3D Generative Radiance Field, Leheng Li, Qing Lian, Luozhou Wang, Ningning Ma, Ying-Cong Chen
33. 3D GAN Inversion With Facial Symmetry Prior, Fei Yin, Yong Zhang, Xuan Wang, Tengfei Wang, Xiaoyu Li, Yuan Gong, Yanbo Fan, Xiaodong Cun, Ying Shan, Cengiz Oztireli, Yujiu Yang
34. StyleIPSB: Identity-Preserving Semantic Basis of StyleGAN for High Fidelity Face Swapping, Diqiong Jiang, Dan Song, Ruofeng Tong, Min Tang
35. FFHQ-UV: Normalized Facial UV-Texture Dataset for 3D Face Reconstruction, Haoran Bai, Di Kang, Haoxian Zhang, Jinshan Pan, Linchao Bao
36. Robust Model-Based Face Reconstruction Through Weakly-Supervised Outlier Segmentation, Chunlu Li, Andreas Morel-Forster, Thomas Vetter, Bernhard Egger, Adam Kortylewski
37. Learning Neural Proto-Face Field for Disentangled 3D Face Modeling in the Wild, Zhenyu Zhang, Renwang Chen, Weijian Cao, Ying Tai, Chengjie Wang
38. A Hierarchical Representation Network for Accurate and Detailed Face Reconstruction From In-the-Wild Images, Biwen Lei, Jianqiang Ren, Mengyang Feng, Miaomiao Cui, Xuansong Xie
39. BlendFields: Few-Shot Example-Driven Facial Modeling, Kacper Kania, Stephan J. Garbin, Andrea Tagliasacchi, Virginia Estellers, Kwang Moo Yi, Julien Valentin, Tomasz Trzcinski, Marek Kowalski
40. Implicit Neural Head Synthesis via Controllable Local Deformation Fields, Chuhan Chen, Matthew O'Toole, Gaurav Bharaj, Pablo Garrido
41. DPE: Disentanglement of Pose and Expression for General Video Portrait Editing, Youxin Pang, Yong Zhang, Weize Quan, Yanbo Fan, Xiaodong Cun, Ying Shan, Dong-Ming Yan
42. GANHead: Towards Generative Animatable Neural Head Avatars, Sijing Wu, Yichao Yan, Yunhao Li, Yuhao Cheng, Wenhan Zhu, Ke Gao, Xiaobo Li, Guangtao Zhai
43. EDGE: Editable Dance Generation From Music, Jonathan Tseng, Rodrigo Castellon, Karen Liu
44. Blowing in the Wind: CycleNet for Human Cinemagraphs From Still Images, Hugo Bertiche, Niloy J. Mitra, Kuldeep Kulkarni, Chun-Hao P. Huang, Tuanfeng Y. Wang, Meysam Madadi, Sergio Escalera, Duygu Ceylan
45. Generating Holistic 3D Human Motion From Speech, Hongwei Yi, Hualin Liang, Yifei Liu, Qiong Cao, Yandong Wen, Timo Bolkart, Dacheng Tao, Michael J. Black
46. Avatars Grow Legs: Generating Smooth Human Motion From Sparse Tracking Inputs With Diffusion Model, Yuming Du, Robin Kips, Albert Pumarola, Sebastian Starke, Ali Thabet, Artsiom Sanakoyeu
47. Learning Anchor Transformations for 3D Garment Animation, Fang Zhao, Zekun Li, Shaoli Huang, Junwu Weng, Tianfei Zhou, Guo-Sen Xie, Jue Wang, Ying Shan
48. CloSET: Modeling Clothed Humans on Continuous Surface With Explicit Template Decomposition, Hongwen Zhang, Siyou Lin, Ruizhi Shao, Yuxiang Zhang, Zerong Zheng, Han Huang, Yandong Guo, Yebin Liu
49. ECON: Explicit Clothed Humans Optimized via Normal ✧ Integration, Yuliang Xiu, Jinlong Yang, Xu Cao, Dimitrios Tzionas, Michael J. Black
50. PersonNeRF: Personalized Reconstruction From Photo Collections, Chung-Yi Weng, Pratul P. Srinivasan, Brian Curless, Ira Kemelmacher-Shlizerman
51. 3D Human Mesh Estimation From Virtual Markers, Xiaoxuan Ma, Jiajun Su, Chunyu Wang, Wentao Zhu, Yizhou Wang
52. Overcoming the Trade-Off Between Accuracy and Plausibility in 3D Hand Shape Reconstruction, Ziwei Yu, Chen Li, Linlin Yang, Xiaoxu Zheng, Michael Bi Mi, Gim Hee Lee, Angela Yao

53. Recovering 3D Hand Mesh Sequence From a Single Blurry Image: A New Dataset and Temporal Unfolding, *Yeonguk Oh, Joonkyu Park, Jaeha Kim, Gyeongsik Moon, Kyoung Mu Lee*
54. MeMaHand: Exploiting Mesh-Mano Interaction for Single Image Two-Hand Reconstruction, *Congyi Wang, Feida Zhu, Shilei Wen*
55. PLIKS: A Pseudo-Linear Inverse Kinematic Solver for 3D Human Body Estimation, *Karthik Shetty, Annette Birkhold, Srikrishna Jaganathan, Norbert Strobel, Markus Kowarschik, Andreas Maier, Bernhard Egger*
56. CAMS: CAnonicalized Manipulation Spaces for Category-Level Functional Hand-Object Manipulation Synthesis, *Juntian Zheng, Qingyuan Zheng, Lixing Fang, Yun Liu, Li Yi*
57. Instant-NVR: Instant Neural Volumetric Rendering for Human-Object Interactions From Monocular RGBD Stream, *Yuheng Jiang, Kaixin Yao, Zhuo Su, Zhehao Shen, Haimin Luo, Lan Xu*
58. BundleSDF: Neural 6-DoF Tracking and 3D Reconstruction of Unknown Objects, *Bowen Wen, Jonathan Tremblay, Valts Blukis, Stephen Tyree, Thomas Müller, Alex Evans, Dieter Fox, Jan Kautz, Stan Birchfield*
59. Human-Art: A Versatile Human-Centric Dataset Bridging Natural and Artificial Scenes, *Xuan Ju, Ailing Zeng, Jianan Wang, Qiang Xu, Lei Zhang*
60. Omnimate3D: Associating Objects and Their Effects in Unconstrained Monocular Video, *Mohammed Suhail, Erika Lu, Zhengqi Li, Noah Snavely, Leonid Sigal, Forrester Cole*
61. On the Benefits of 3D Pose and Tracking for Human Action Recognition, *Jathushan Rajasegaran, Georgios Pavlakos, Angjoo Kanazawa, Christoph Feichtenhofer, Jitendra Malik*
62. Towards Stable Human Pose Estimation via Cross-View Fusion and Foot Stabilization, *Li'an Zhuo, Jian Cao, Qi Wang, Bang Zhang, Liefeng Bo*
63. Human Pose As Compositional Tokens, *Zigang Geng, Chunyu Wang, Yixuan Wei, Ze Liu, Houqiang Li, Han Hu*
64. PoseExaminer: Automated Testing of Out-of-Distribution Robustness in Human Pose and Shape Estimation, *Qihao Liu, Adam Kortylewski, Alan L. Yuille*
65. SLOPER4D: A Scene-Aware Dataset for Global 4D Human Pose Estimation in Urban Environments, *Yudi Dai, Yitai Lin, Xiping Lin, Chenglu Wen, Lan Xu, Hongwei Yi, Siqi Shen, Yuexin Ma, Cheng Wang*
66. Semi-Supervised 2D Human Pose Estimation Driven by Position Inconsistency Pseudo Label Correction Module, *Linzhi Huang, Yulong Li, Hongbo Tian, Yue Yang, Xiangang Li, Weihong Deng, Jieping Ye*
67. Human Pose Estimation in Extremely Low-Light Conditions, *Sohyun Lee, Jaesung Rim, Boseung Jeong, Geonu Kim, Byungju Woo, Haechan Lee, Sunghyun Cho, Suha Kwak*
68. Flexible- C^m GAN: Towards Precise 3D Dose Prediction in Radiotherapy, *Riqiang Gao, Bin Lou, Zhoubing Xu, Dorin Comaniciu, Ali Kamen*
69. DualRefine: Self-Supervised Depth and Pose Estimation Through Iterative Epipolar Sampling and Refinement Toward Equilibrium, *Antyanta Bangunharcana, Ahmed Magd, Kyung-Soo Kim*
70. A Rotation-Translation-Decoupled Solution for Robust and Efficient Visual-Inertial Initialization, *Yijia He, Bo Xu, Zhanpeng Ouyang, Hongdong Li*
71. Semidefinite Relaxations for Robust Multiview Triangulation, *Linus Härenstam-Nielsen, Niclas Zeller, Daniel Cremers*
72. A Probabilistic Attention Model With Occlusion-Aware Texture Regression for 3D Hand Reconstruction From a Single RGB Image, *Zheheng Jiang, Hossein Rahmani, Sue Black, Bryan M. Williams*
73. Instant Multi-View Head Capture Through Learnable Registration, *Timo Bolkart, Tianye Li, Michael J. Black*
74. On the Importance of Accurate Geometry Data for Dense 3D Vision Tasks, *HyunJun Jung, Patrick Ruhkamp, Guangyao Zhai, Nikolas Brasch, Yitong Li, Yannick Verdie, Jifei Song, Yiren Zhou, Anil Armagan, Slobodan Ilic, Aleš Leonardis, Nassir Navab, Benjamin Busam*
75. Learning 3D Scene Priors With 2D Supervision, *Yinyu Nie, Angela Dai, Xiaoguang Han, Matthias Nießner*
76. OmniObject3D: Large-Vocabulary 3D Object Dataset for Realistic Perception, Reconstruction and Generation, *Tong Wu, Jiarui Zhang, Xiao Fu, Yuxin Wang, Jiawei Ren, Liang Pan, Wayne Wu, Lei Yang, Jiaqi Wang, Chen Qian, Dahua Lin, Ziwei Liu*
77. OpenScene: 3D Scene Understanding With Open Vocabularies, *Songyou Peng, Kyle Genova, Chiyu "Max" Jiang, Andrea Tagliasacchi, Marc Pollefeys, Thomas Funkhouser*
78. Multi-View Azimuth Stereo via Tangent Space Consistency, *Xu Cao, Hiroaki Santo, Fumio Okura, Yasuyuki Matsushita*
79. Progressive Transformation Learning for Leveraging Virtual Images in Training, *Yi-Ting Shen, Hyungtae Lee, Heesung Kwon, Shuvra S. Bhattacharyya*
80. Connecting the Dots: Floorplan Reconstruction Using Two-Level Queries, *Yuanwen Yue, Theodora Kontogianni, Konrad Schindler, Francis Engelmann*
81. NeRF-Supervised Deep Stereo, *Fabio Tosi, Alessio Tonioni, Daniele De Gregorio, Matteo Poggi*
82. Semantic Scene Completion With Cleaner Self, *Fengyun Wang, Dong Zhang, Hanwang Zhang, Jinhui Tang, Qianru Sun*
83. PanelNet: Understanding 360 Indoor Environment via Panel Representation, *Haozheng Yu, Lu He, Bing Jian, Weiwei Feng, Shan Liu*
84. Implicit View-Time Interpolation of Stereo Videos Using Multi-Plane Disparities and Non-Uniform Coordinates, *Avinash Paliwal, Andrii Tsarov, Nima Khademi Kalantari*
85. Depth Estimation From Indoor Panoramas With Neural Scene Representation, *Wenjie Chang, Yueyi Zhang, Zhiwei Xiong*
86. NeuralPCI: Spatio-Temporal Neural Field for 3D Point Cloud Multi-Frame Non-Linear Interpolation, *Zehan Zheng, Danni Wu, Ruisi Lu, Fan Lu, Guang Chen, Changjun Jiang*
87. RIAV-MVS: Recurrent-Indexing an Asymmetric Volume for Multi-View Stereo, *Changjiang Cai, Pan Ji, Qingan Yan, Yi Xu*
88. NeuMap: Neural Coordinate Mapping by Auto-Transdecoder for Camera Localization, *Shitao Tang, Sicong Tang, Andrea Tagliasacchi, Ping Tan, Yasutaka Furukawa*
89. MACARONS: Mapping and Coverage Anticipation With RGB Online Self-Supervision, *Antoine Guédon, Tom Monnier, Pascal Monasse, Vincent Lepetit*
90. vMAP: Vectorised Object Mapping for Neural Field SLAM, *Xin Kong, Shikun Liu, Marwan Taher, Andrew J. Davison*
91. Seeing a Rose in Five Thousand Ways, *Yunzhi Zhang, Shangzhe Wu, Noah Snavely, Jiajun Wu*
92. Propagate and Calibrate: Real-Time Passive Non-Line-of-Sight Tracking, *Yihao Wang, Zhigang Wang, Bin Zhao, Dong Wang, Mulin Chen, Xuelong Li*

93. Seeing With Sound: Long-range Acoustic Beamforming for Multimodal Scene Understanding, *Praneeth Chakravarthula, Jim Aldon D'Souza, Ethan Tseng, Joe Bartusek, Felix Heide*
94. Distilling Focal Knowledge From Imperfect Expert for 3D Object Detection, *Jia Zeng, Li Chen, Hanming Deng, Lewei Lu, Junchi Yan, Yu Qiao, Hongyang Li*
95. BEV-LaneDet: An Efficient 3D Lane Detection Based on Virtual Camera via Key-Points, *Ruihao Wang, Jian Qin, Kaiying Li, Yaochen Li, Dong Cao, Jintao Xu*
96. AShapeFormer: Semantics-Guided Object-Level Active Shape Encoding for 3D Object Detection via Transformers, *Zechuan Li, Hongshan Yu, Zhengeng Yang, Tongjia Chen, Naveed Akhtar*
97. Benchmarking Robustness of 3D Object Detection to Common Corruptions, *Yinpeng Dong, Caixin Kang, Jinlai Zhang, Zijian Zhu, Yikai Wang, Xiao Yang, Hang Su, Xingxing Wei, Jun Zhu*
98. Gaussian Label Distribution Learning for Spherical Image Object Detection, *Hang Xu, Xinyuan Liu, Qiang Zhao, Yike Ma, Chenggang Yan, Feng Dai*
99. Deep Depth Estimation From Thermal Image, *Ukcheol Shin, Jinsun Park, In So Kweon*
100. LidarGait: Benchmarking 3D Gait Recognition With Point Clouds, *Chuanfu Shen, Chao Fan, Wei Wu, Rui Wang, George Q. Huang, Shiqi Yu*
101. Generalized UAV Object Detection via Frequency Domain Disentanglement, *Kunyu Wang, Xueyang Fu, Yukun Huang, Chengzhi Cao, Gege Shi, Zheng-Jun Zha*
102. Learning Compact Representations for LiDAR Completion and Generation, *Yuwen Xiong, Wei-Chiu Ma, Jingkang Wang, Raquel Urtasun*
103. CXTrack: Improving 3D Point Cloud Tracking With Contextual Information, *Tian-Xing Xu, Yuan-Chen Guo, Yu-Kun Lai, Song-Hai Zhang*
104. Multispectral Video Semantic Segmentation: A Benchmark Dataset and Baseline, *Wei Ji, Jingjing Li, Cheng Bian, Zongwei Zhou, Jiaying Zhao, Alan L. Yuille, Li Cheng*
105. LinK: Linear Kernel for LiDAR-Based 3D Perception, *Tao Lu, Xiang Ding, Haisong Liu, Gangshan Wu, Limin Wang*
106. Point Cloud Forecasting as a Proxy for 4D Occupancy Forecasting, *Tarasha Khurana, Peiyun Hu, David Held, Deva Ramanan*
107. Curricular Object Manipulation in LiDAR-Based Object Detection, *Ziyue Zhu, Qiang Meng, Xiao Wang, Ke Wang, Liujiang Yan, Jian Yang*
108. Delivering Arbitrary-Modal Semantic Segmentation, *Jiaming Zhang, Ruiping Liu, Hao Shi, Kailun Yang, Simon Reiß, Kunyu Peng, Haodong Fu, Kaiwei Wang, Rainer Stiefelhofen*
109. Robust Outlier Rejection for 3D Registration With Variational Bayes, *Haobo Jiang, Zheng Dang, Zhen Wei, Jin Xie, Jian Yang, Mathieu Salzmann*
110. 3D Human Keypoints Estimation From Point Clouds in the Wild Without Human Labels, *Zhenzhen Weng, Alexander S. Gorban, Jingwei Ji, Mahyar Najibi, Yin Zhou, Dragomir Anguelov*
111. Self-Supervised Pre-Training With Masked Shape Prediction for 3D Scene Understanding, *Li Jiang, Zetong Yang, Shaoshuai Shi, Vladislav Golyanik, Dengxin Dai, Bernt Schiele*
112. ULIP: Learning a Unified Representation of Language, Images, and Point Clouds for 3D Understanding, *Le Xue, Mingfei Gao, Chen Xing, Roberto Martín-Martín, Jiajun Wu, Caiming Xiong, Ran Xu, Juan Carlos Niebles, Silvio Savarese*
113. Open-Vocabulary Point-Cloud Object Detection Without 3D Annotation, *Yuheng Lu, Chenfeng Xu, Xiaobao Wei, Xiaodong Xie, Masayoshi Tomizuka, Kurt Keutzer, Shanghang Zhang*
114. FlatFormer: Flattened Window Attention for Efficient Point Cloud Transformer, *Zhijian Liu, Xinyu Yang, Haotian Tang, Shang Yang, Song Han*
115. PointCMP: Contrastive Mask Prediction for Self-Supervised Learning on Point Cloud Videos, *Zhiqiang Shen, Xiaoxiao Sheng, Longguang Wang, Yulan Guo, Qiong Liu, Xi Zhou*
116. E2PN: Efficient SE(3)-Equivariant Point Network, *Minghan Zhu, Maani Ghaffari, William A. Clark, Huei Peng*
117. Poly-PC: A Polyhedral Network for Multiple Point Cloud Tasks at Once, *Tao Xie, Shiguang Wang, Ke Wang, Linqi Yang, Zhiqiang Jiang, Xingcheng Zhang, Kun Dai, Ruifeng Li, Jian Cheng*
118. Improving Graph Representation for Point Cloud Segmentation via Attentive Filtering, *Nan Zhang, Zhiyi Pan, Thomas H. Li, Wei Gao, Ge Li*
119. BUFFER: Balancing Accuracy, Efficiency, and Generalizability in Point Cloud Registration, *Sheng Ao, Qingyong Hu, Hanyun Wang, Kai Xu, Yulan Guo*
120. TopDiG: Class-Agnostic Topological Directional Graph Extraction From Remote Sensing Images, *Bingnan Yang, Mi Zhang, Zhan Zhang, Zhili Zhang, Xiangyun Hu*
121. Recognizing Rigid Patterns of Unlabeled Point Clouds by Complete and Continuous Isometry Invariants With No False Negatives and No False Positives, *Daniel Widdowson, Vitaliy Kurlin*
122. Both Style and Distortion Matter: Dual-Path Unsupervised Domain Adaptation for Panoramic Semantic Segmentation, *Xu Zheng, Jinjing Zhu, Yexin Liu, Zidong Cao, Chong Fu, Lin Wang*
123. CCuantuMM: Cycle-Consistent Quantum-Hybrid Matching of Multiple Shapes, *Harshil Bhatia, Edith Tretschk, Zorah Löhner, Marcel Seelbach Benkner, Michael Moeller, Christian Theobalt, Vladislav Golyanik*
124. Enhancing Deformable Local Features by Jointly Learning to Detect and Describe Keypoints, *Guilherme Potje, Felipe Cadar, André Araujo, Renato Martins, Erickson R. Nascimento*
125. Understanding and Improving Features Learned in Deep ✨ Functional Maps, *Souhaib Attaiki, Maks Ovsjanikov*
126. High-Frequency Stereo Matching Network, *Haoliang Zhao, ✨ Huizhou Zhou, Yongjun Zhang, Jie Chen, Yitong Yang, Yong Zhao*
127. Rethinking Optical Flow From Geometric Matching Consistent Perspective, *Qiaole Dong, Chenjie Cao, Yanwei Fu*
128. Efficient Robust Principal Component Analysis via Block Krylov Iteration and CUR Decomposition, *Shun Fang, Zhengqin Xu, Shiqian Wu, Shoulie Xie*
129. VectorFloorSeg: Two-Stream Graph Attention Network for ✨ Vectorized Roughcast Floorplan Segmentation, *Bingchen Yang, Haiyong Jiang, Hao Pan, Jun Xiao*
130. TBP-Former: Learning Temporal Bird's-Eye-View Pyramid for Joint Perception and Prediction in Vision-Centric Autonomous Driving, *Shaoheng Fang, Zi Wang, Yiqi Zhong, Junhao Ge, Siheng Chen*
131. Implicit Occupancy Flow Fields for Perception and Prediction in ✨ Self-Driving, *Ben Agro, Quinlan Sykora, Sergio Casas, Raquel Urtasun*
132. UniSim: A Neural Closed-Loop Sensor Simulator, *Ze Yang, Yun ✨ Chen, Jingkang Wang, Sivabalan Manivasagam, Wei-Chiu Ma, Anqi Joyce Yang, Raquel Urtasun*

133. FEND: A Future Enhanced Distribution-Aware Contrastive Learning Framework for Long-Tail Trajectory Prediction, *Yuning Wang, Pu Zhang, Lei Bai, Jianru Xue*
134. EqMotion: Equivariant Multi-Agent Motion Prediction With Invariant Interaction Reasoning, *Chenxin Xu, Robby T. Tan, Yuhong Tan, Siheng Chen, Yu Guang Wang, Xinchao Wang, Yanfeng Wang*
135. Lookahead Diffusion Probabilistic Models for Refining Mean Estimation, *Guoqiang Zhang, Kenta Niwa, W. Bastiaan Kleijn*
136. Neural Volumetric Memory for Visual Locomotion Control, ✧ *Ruihan Yang, Ge Yang, Xiaolong Wang*
137. Gazeformer: Scalable, Effective and Fast Prediction of Goal-Directed Human Attention, *Sounak Mondal, Zhibo Yang, Seoyoung Ahn, Dimitris Samaras, Gregory Zelinsky, Minh Hoai*
138. DrapeNet: Garment Generation and Self-Supervised Draping, *Luca De Luigi, Ren Li, Benoît Guillard, Mathieu Salzmann, Pascal Fua*
139. Tracking Multiple Deformable Objects in Egocentric Videos, *Mingzhen Huang, Xiaoxing Li, Jun Hu, Honghong Peng, Siwei Lyu*
140. Good Is Bad: Causality Inspired Cloth-Debiasing for Cloth-Changing Person Re-Identification, *Zhengwei Yang, Meng Lin, Xian Zhong, Yu Wu, Zheng Wang*
141. Micron-BERT: BERT-Based Facial Micro-Expression Recognition, *Xuan-Bac Nguyen, Chi Nhan Duong, Xin Li, Susan Gauch, Han-Seok Seo, Khoa Luu*
142. MARLIN: Masked Autoencoder for Facial Video Representation LearnINg, *Zhixi Cai, Shreya Ghosh, Kalin Stefanov, Abhinav Dhall, Jianfei Cai, Hamid Reza Tofighi, Reza Haffari, Munawar Hayat*
143. StyleSync: High-Fidelity Generalized and Personalized Lip Sync in Style-Based Generator, *Jiazhi Guan, Zhanwang Zhang, Hang Zhou, Tianshu Hu, Kaisiyuan Wang, Dongliang He, Haocheng Feng, Jingtuo Liu, Errui Ding, Ziwei Liu, Jingdong Wang*
144. REALIMPACT: A Dataset of Impact Sound Fields for Real Objects, ✧ *Samuel Clarke, Ruohan Gao, Mason Wang, Mark Rau, Julia Xu, Jui-Hsien Wang, Doug L. James, Jiajun Wu*
145. STMT: A Spatial-Temporal Mesh Transformer for MoCap-Based Action Recognition, *Xiaoyu Zhu, Po-Yao Huang, Junwei Liang, Celso M. de Melo, Alexander G. Hauptmann*
146. Progressive Spatio-Temporal Alignment for Efficient Event-Based Motion Estimation, *Xueyan Huang, Yueyi Zhang, Zhiwei Xiong*
147. Event-Based Shape From Polarization, *Manasi Muglikar, Leonard Bayersfeld, Diederik Paul Moeys, Davide Scaramuzza*
148. Learning Spatial-Temporal Implicit Neural Representations for Event-Guided Video Super-Resolution, *Yunfan Lu, Zipeng Wang, Minjie Liu, Hongjian Wang, Lin Wang*
149. BiFormer: Learning Bilateral Motion Estimation via Bilateral Transformer for 4K Video Frame Interpolation, *Junheum Park, Jintae Kim, Chang-Su Kim*
150. A Unified Pyramid Recurrent Network for Video Frame Interpolation, *Xin Jin, Longhai Wu, Jie Chen, Youxin Chen, Jayoon Koo, Cheul-hee Hahm*
151. Event-Based Blurry Frame Interpolation Under Blind Exposure, *Wenming Weng, Yueyi Zhang, Zhiwei Xiong*
152. FlowFormer++: Masked Cost Volume Autoencoding for Pretraining Optical Flow Estimation, *Xiaoyu Shi, Zhaoyang Huang, Dasong Li, Manyuan Zhang, Ka Chun Cheung, Simon See, Hongwei Qin, Jifeng Dai, Hongsheng Li*
153. POTTER: Pooling Attention Transformer for Efficient Human Mesh Recovery, *Ce Zheng, Xianpeng Liu, Guo-Jun Qi, Chen Chen*
154. Adaptive Patch Deformation for Textureless-Resilient Multi-View Stereo, *Yuesong Wang, Zhaojie Zeng, Tao Guan, Wei Yang, Zhuo Chen, Wenkai Liu, Luoyuan Xu, Yawei Luo*
155. On the Difficulty of Unpaired Infrared-to-Visible Video Translation: Fine-Grained Content-Rich Patches Transfer, *Zhenjie Yu, Shuang Li, Yirui Shen, Chi Harold Liu, Shuigen Wang*
156. Thermal Spread Functions (TSF): Physics-Guided Material Classification, *Aniket Dashpute, Vishwanath Saragadam, Emma Alexander, Florian Willomitzer, Aggelos Katsaggelos, Ashok Veeraraghavan, Oliver Cossairt*
157. Better "CMOS" Produces Clearer Images: Learning Space-Variant Blur Estimation for Blind Image Super-Resolution, *Xuhai Chen, Jiangning Zhang, Chao Xu, Yabiao Wang, Chengjie Wang, Yong Liu*
158. Learning Semantic-Aware Knowledge Guidance for Low-Light Image Enhancement, *Yuhui Wu, Chen Pan, Guoqing Wang, Yang Yang, Jiwei Wei, Chongyi Li, Heng Tao Shen*
159. CutMIB: Boosting Light Field Super-Resolution via Multi-View Image Blending, *Zeyu Xiao, Yutong Liu, Ruisheng Gao, Zhiwei Xiong*
160. sRGB Real Noise Synthesizing With Neighboring Correlation-Aware Noise Model, *Zixuan Fu, Lanqing Guo, Bihan Wen*
161. Masked Image Training for Generalizable Deep Image Denoising, *Haoyu Chen, Jinjin Gu, Yihao Liu, Salma Abdel Magid, Chao Dong, Qiong Wang, Hanspeter Pfister, Lei Zhu*
162. DR2: Diffusion-Based Robust Degradation Remover for Blind Face Restoration, *Zhixin Wang, Ziyang Zhang, Xiaoyun Zhang, Huangjie Zheng, Mingyuan Zhou, Ya Zhang, Yanfeng Wang*
163. Learning Distortion Invariant Representation for Image Restoration From a Causality Perspective, *Xin Li, Bingchen Li, Xin Jin, Cuiling Lan, Zhibo Chen*
164. Perception-Oriented Single Image Super-Resolution Using Optimal Objective Estimation, *Seung Ho Park, Young Su Moon, Nam Ik Cho*
165. Catch Missing Details: Image Reconstruction With Frequency Augmented Variational Autoencoder, *Xinmiao Lin, Yikang Li, Jenhao Hsiao, Chiuman Ho, Yu Kong*
166. MD-VQA: Multi-Dimensional Quality Assessment for UGC Live Videos, *Zicheng Zhang, Wei Wu, Wei Sun, Danyang Tu, Wei Lu, Xiongkuo Min, Ying Chen, Guangtao Zhai*
167. CABM: Content-Aware Bit Mapping for Single Image Super-Resolution Network With Large Input, *Senmao Tian, Ming Lu, Jiaming Liu, Yandong Guo, Yurong Chen, Shunli Zhang*
168. Initialization Noise in Image Gradients and Saliency Maps, *Ann-Christin Woerl, Jan Disselhoff, Michael Wand*
169. Local Implicit Normalizing Flow for Arbitrary-Scale Image Super-Resolution, *Jie-En Yao, Li-Yuan Tsao, Yi-Chen Lo, Roy Tseng, Chia-Che Chang, Chun-Yi Lee*
170. Deep Arbitrary-Scale Image Super-Resolution via Scale-Equivariance Pursuit, *Xiaohang Wang, Xuanhong Chen, Bingbing Ni, Hang Wang, Zhengyan Tong, Yutian Liu*
171. CiaoSR: Continuous Implicit Attention-in-Attention Network for Arbitrary-Scale Image Super-Resolution, *Jie Zhang Cao, Qin Wang, Yongqin Xian, Yawei Li, Bingbing Ni, Zhiming Pi, Kai Zhang, Yulun Zhang, Radu Timofte, Luc Van Gool*
172. Multiplicative Fourier Level of Detail, *Yishun Dou, Zhong Zheng, ✧ Qiaoqiao Jin, Bingbing Ni*

173. Document Image Shadow Removal Guided by Color-Aware Background, *Ling Zhang, Yinghao He, Qing Zhang, Zheng Liu, Xiaolong Zhang, Chunxia Xiao*
174. StyleRes: Transforming the Residuals for Real Image Editing With StyleGAN, *Hamza Pehlivan, Yusuf Dalva, Aysegul Dundar*
175. TopNet: Transformer-Based Object Placement Network for Image Compositing, *Sijie Zhu, Zhe Lin, Scott Cohen, Jason Kuen, Zhifei Zhang, Chen Chen*
176. VecFontSDF: Learning to Reconstruct and Synthesize High-Quality Vector Fonts via Signed Distance Functions, *Zeqing Xia, Bojun Xiong, Zhouhui Lian*
177. CF-Font: Content Fusion for Few-Shot Font Generation, *Chi Wang, Min Zhou, Tiezheng Ge, Yuning Jiang, Hujun Bao, Weiwei Xu*
178. SIEDOB: Semantic Image Editing by Disentangling Object and ✧ Background, *Wuyang Luo, Su Yang, Xinjian Zhang, Weishan Zhang*
179. MaskSketch: Unpaired Structure-Guided Masked Image ✧ Generation, *Dina Bashkirova, José Lezama, Kihyuk Sohn, Kate Saenko, Ifan Essa*
180. Text2Scene: Text-Driven Indoor Scene Stylization With Part-Aware Details, *Inwoo Hwang, Hyeonwoo Kim, Young Min Kim*
181. Uncovering the Disentanglement Capability in Text-to-Image Diffusion Models, *Qiucheng Wu, Yujian Liu, Handong Zhao, Ajinkya Kale, Trung Bui, Tong Yu, Zhe Lin, Yang Zhang, Shiyu Chang*
182. VectorFusion: Text-to-SVG by Abstracting Pixel-Based Diffusion Models, *Ajay Jain, Amber Xie, Pieter Abbeel*
183. Plug-and-Play Diffusion Features for Text-Driven Image-to-Image Translation, *Narek Tumanyan, Michal Geyer, Shai Bagon, Tali Dekel*
184. Multi-Concept Customization of Text-to-Image Diffusion, *Nupur Kumari, Bingliang Zhang, Richard Zhang, Eli Shechtman, Jun-Yan Zhu*
185. Unifying Layout Generation With a Decoupled Diffusion Model, *Mude Hui, Zhizheng Zhang, Xiaoyi Zhang, Wenxuan Xie, Yuwang Wang, Yan Lu*
186. BBDM: Image-to-Image Translation With Brownian Bridge Diffusion Models, *Bo Li, Kaitao Xue, Bin Liu, Yu-Kun Lai*
187. Towards Practical Plug-and-Play Diffusion Models, *Hyojun Go, Yunsung Lee, Jin-Young Kim, Seunghyun Lee, Myeongho Jeong, Hyun Seung Lee, Seungtaek Choi*
188. Post-Training Quantization on Diffusion Models, *Yuzhang Shang, Zhihang Yuan, Bin Xie, Bingzhe Wu, Yan Yan*
189. DiffTalk: Crafting Diffusion Models for Generalized Audio-Driven Portraits Animation, *Shuai Shen, Wenliang Zhao, Zibin Meng, Wanhua Li, Zheng Zhu, Jie Zhou, Jiwen Lu*
190. Mask-Guided Matting in the Wild, *Kwanyong Park, Sanghyun Woo, Seoung Wug Oh, In So Kweon, Joon-Young Lee*
191. Not All Image Regions Matter: Masked Vector Quantization for Autoregressive Image Generation, *Mengqi Huang, Zhendong Mao, Quan Wang, Yongdong Zhang*
192. Compression-Aware Video Super-Resolution, *Yingwei Wang, Takashi Isobe, Xu Jia, Xin Tao, Huchuan Lu, Yu-Wing Tai*
193. Neural Rate Estimator and Unsupervised Learning for Efficient Distributed Image Analytics in Split-DNN Models, *Nilesh Ahuja, Parul Datta, Bhavya Kanzariya, V. Srinivasa Somayazulu, Omesh Tickoo*
194. DNeRV: Modeling Inherent Dynamics via Difference Neural Representation for Videos, *Qi Zhao, M. Salman Asif, Zhan Ma*
195. Polynomial Implicit Neural Representations for Large Diverse ✧ Datasets, *Rajhans Singh, Ankita Shukla, Pavan Turaga*
196. Learning Decorrelated Representations Efficiently Using Fast Fourier Transform, *Yutaro Shigeto, Masashi Shimbo, Yuya Yoshikawa, Akikazu Takeuchi*
197. SparseViT: Revisiting Activation Sparsity for Efficient High-Resolution Vision Transformer, *Xuanyao Chen, Zhijian Liu, Haotian Tang, Li Yi, Hang Zhao, Song Han*
198. N-Gram in Swin Transformers for Efficient Lightweight Image Super-Resolution, *Haram Choi, Jeongmin Lee, Jihoon Yang*
199. Slide-Transformer: Hierarchical Vision Transformer With Local Self-Attention, *Xuran Pan, Tianzhu Ye, Zhuofan Xia, Shiji Song, Gao Huang*
200. Joint Token Pruning and Squeezing Towards More Aggressive Compression of Vision Transformers, *Siyuan Wei, Tianzhu Ye, Shen Zhang, Yao Tang, Jiajun Liang*
201. Top-Down Visual Attention From Analysis by Synthesis, *Baifeng ✧ Shi, Trevor Darrell, Xin Wang*
202. Probing Neural Representations of Scene Perception in a Hippocampally Dependent Task Using Artificial Neural Networks, *Markus Frey, Christian F. Doeller, Caswell Barry*
203. Masked Image Modeling With Local Multi-Scale Reconstruction, ✧ *Haoqing Wang, Yehui Tang, Yunhe Wang, Jianyuan Guo, Zhi-Hong Deng, Kai Han*
204. Siamese Image Modeling for Self-Supervised Vision Representation Learning, *Chenxin Tao, Xizhou Zhu, Weijie Su, Gao Huang, Bin Li, Jie Zhou, Yu Qiao, Xiaogang Wang, Jifeng Dai*
205. MAGE: MAsked Generative Encoder to Unify Representation Learning and Image Synthesis, *Tianhong Li, Huiwen Chang, Shlok Mishra, Han Zhang, Dina Katabi, Dilip Krishnan*
206. Diverse Embedding Expansion Network and Low-Light Cross-Modality Benchmark for Visible-Infrared Person Re-Identification, *Yukang Zhang, Hanzi Wang*
207. DistilPose: Tokenized Pose Regression With Heatmap Distillation, *Suhang Ye, Yingyi Zhang, Jie Hu, Liujuan Cao, Shengchuan Zhang, Lei Shen, Jun Wang, Shouhong Ding, Rongrong Ji*
208. Graph Transformer GANs for Graph-Constrained House Generation, *Hao Tang, Zhenyu Zhang, Humphrey Shi, Bo Li, Ling Shao, Nicu Sebe, Radu Timofte, Luc Van Gool*
209. Automatic High Resolution Wire Segmentation and Removal, *Mang Tik Chiu, Xuaner Zhang, Zijun Wei, Yuqian Zhou, Eli Shechtman, Connelly Barnes, Zhe Lin, Florian Kainz, Sohrab Amirghodsi, Humphrey Shi*
210. Tree Instance Segmentation With Temporal Contour Graph, *Adnan Firoze, Cameron Wingren, Raymond A. Yeh, Bedrich Benes, Daniel Aliaga*
211. Dual-Path Adaptation From Image to Video Transformers, *Jungin Park, Jiyoung Lee, Kwanghoon Sohn*
212. Rethinking Video ViTs: Sparse Video Tubes for Joint Image and Video Learning, *AJ Piergiovanni, Weicheng Kuo, Anelia Angelova*
213. Modeling Video As Stochastic Processes for Fine-Grained Video ✧ Representation Learning, *Heng Zhang, Daqing Liu, Qi Zheng, Bing Su*
214. Masked Motion Encoding for Self-Supervised Video Representation Learning, *Xinyu Sun, Peihao Chen, Liangwei Chen, Changhao Li, Thomas H. Li, Minghui Tan, Chuang Gan*

215. Boosting Video Object Segmentation via Space-Time Correspondence Learning, *Yurong Zhang, Liulei Li, Wenguan Wang, Rong Xie, Li Song, Wenjun Zhang*
216. Two-Shot Video Object Segmentation, *Kun Yan, Xiao Li, Fangyun Wei, Jinglu Wang, Chenbin Zhang, Ping Wang, Yan Lu*
217. Look Before You Match: Instance Understanding Matters in Video Object Segmentation, *Junke Wang, Dongdong Chen, Zuxuan Wu, Chong Luo, Chuanxin Tang, Xiyang Dai, Yucheng Zhao, Yujia Xie, Lu Yuan, Yu-Gang Jiang*
218. Spatial-Then-Temporal Self-Supervised Learning for Video Correspondence, *Rui Li, Dong Liu*
219. Few-Shot Referring Relationships in Videos, *Yogesh Kumar, Anand Mishra*
220. Vision Transformers Are Parameter-Efficient Audio-Visual Learners, *Yan-Bo Lin, Yi-Lin Sung, Jie Lei, Mohit Bansal, Gedas Bertasius*
221. Egocentric Video Task Translation, *Zihui Xue, Yale Song, Kristen Grauman, Lorenzo Torresani*
222. QPGesture: Quantization-Based and Phase-Guided Motion Matching for Natural Speech-Driven Gesture Generation, *Sicheng Yang, Zhiyong Wu, Minglei Li, Zhensong Zhang, Lei Hao, Weihong Bao, Haolin Zhuang*
223. Co-Speech Gesture Synthesis by Reinforcement Learning With Contrastive Pre-Trained Rewards, *Mingyang Sun, Mengchen Zhao, Yaqing Hou, Minglei Li, Huang Xu, Songcen Xu, Jianye Hao*
224. TimeBalance: Temporally-Invariant and Temporally-Distinctive Video Representations for Semi-Supervised Action Recognition, *Ishan Rajendrakumar Dave, Mamshad Nayeem Rizve, Chen Chen, Mubarak Shah*
225. How Can Objects Help Action Recognition? *Xingyi Zhou, Anurag Arnab, Chen Sun, Cordelia Schmid*
226. Actionlet-Dependent Contrastive Learning for Unsupervised Skeleton-Based Action Recognition, *Lilang Lin, Jiahang Zhang, Jiaying Liu*
227. Decomposed Cross-Modal Distillation for RGB-Based Temporal Action Detection, *Pilhyeon Lee, Taeh Kim, Minho Shim, Dongyoon Wee, Hyeran Byun*
228. ASPnet: Action Segmentation With Shared-Private Representation of Multiple Data Sources, *Beatrice van Amsterdam, Abdolrahim Kadkhodamohammadi, Imanol Luengo, Danail Stoyanov*
229. Proposal-Based Multiple Instance Learning for Weakly-Supervised Temporal Action Localization, *Huan Ren, Wenfei Yang, Tianzhu Zhang, Yongdong Zhang*
230. LOGO: A Long-Form Video Dataset for Group Action Quality Assessment, *Shiyi Zhang, Wenxun Dai, Sujia Wang, Xiangwei Shen, Jiwen Lu, Jie Zhou, Yansong Tang*
231. Use Your Head: Improving Long-Tail Video Recognition, *Toby Perrett, Saptarshi Sinha, Tilo Burghardt, Majid Mirmehdi, Dima Damen*
232. Conditional Generation of Audio From Video via Foley Analogies, *Yuexi Du, Ziyang Chen, Justin Salamon, Bryan Russell, Andrew Owens*
233. Weakly Supervised Video Representation Learning With Unaligned Text for Sequential Videos, *Sixun Dong, Huazhang Hu, Dongze Lian, Weixin Luo, Yicheng Qian, Shenghua Gao*
234. You Can Ground Earlier Than See: An Effective and Efficient Pipeline for Temporal Sentence Grounding in Compressed Videos, *Xiang Fang, Daizong Liu, Pan Zhou, Guoshun Nan*
235. Connecting Vision and Language With Video Localized Narratives, *Paul Voigtlaender, Soravit Changpinyo, Jordi Pont-Tuset, Radu Soricut, Vittorio Ferrari*
236. Video-Text As Game Players: Hierarchical Banzhaf Interaction for Cross-Modal Representation Learning, *Peng Jin, Jinfa Huang, Pengfei Xiong, Shangxuan Tian, Chang Liu, Xiangyang Ji, Li Yuan, Jie Chen*
237. Aligning Step-by-Step Instructional Diagrams to Video Demonstrations, *Jiahao Zhang, Anoop Cherian, Yanbin Liu, Yizhak Ben-Shabat, Cristian Rodriguez, Stephen Gould*
238. Make-a-Story: Visual Memory Conditioned Consistent Story Generation, *Tanzila Rahman, Hsin-Ying Lee, Jian Ren, Sergey Tulyakov, Shweta Mahajan, Leonid Sigal*
239. Test of Time: Instilling Video-Language Models With a Sense of Time, *Piyush Bagad, Makarand Tapaswi, Cees G. M. Snoek*
240. How You Feelin'? Learning Emotions and Mental States in Movie Scenes, *Dhruv Srivastava, Aditya Kumar Singh, Makarand Tapaswi*
241. Continuous Sign Language Recognition With Correlation Network, *Lianyu Hu, Liqing Gao, Zekang Liu, Wei Feng*
242. DIP: Dual Incongruity Perceiving Network for Sarcasm Detection, *Changsong Wen, Guoli Jia, Jufeng Yang*
243. Gloss Attention for Gloss-Free Sign Language Translation, *Aoxiong Yin, Tianyun Zhong, Li Tang, Weike Jin, Tao Jin, Zhou Zhao*
244. Object-Goal Visual Navigation via Effective Exploration of Relations Among Historical Navigation States, *Heming Du, Lincheng Li, Zi Huang, Xin Yu*
245. Behavioral Analysis of Vision-and-Language Navigation Agents, *Zijiao Yang, Arjun Majumdar, Stefan Lee*
246. KERM: Knowledge Enhanced Reasoning for Vision-and-Language Navigation, *Xiangyang Li, Zihan Wang, Jiahao Yang, Yaowei Wang, Shuqiang Jiang*
247. Where Is My Wallet? Modeling Object Proposal Sets for Egocentric Visual Query Localization, *Mengmeng Xu, Yanghao Li, Cheng-Yang Fu, Bernard Ghanem, Tao Xiang, Juan-Manuel Pérez-Rúa*
248. Efficient Multimodal Fusion via Interactive Prompting, *Yaowei Li, Ruijie Quan, Linchao Zhu, Yi Yang*
249. NS3D: Neuro-Symbolic Grounding of 3D Objects and Relations, *Joy Hsu, Jiayuan Mao, Jiajun Wu*
250. Dynamic Inference With Grounding Based Vision and Language Models, *Burak Uzkent, Amanmeet Garg, Wentao Zhu, Keval Doshi, Jingru Yi, Xiaolong Wang, Mohamed Omar*
251. Improving Commonsense in Vision-Language Models via Knowledge Graph Riddles, *Shuquan Ye, Yujia Xie, Dongdong Chen, Yichong Xu, Lu Yuan, Chenguang Zhu, Jing Liao*
252. S3C: Semi-Supervised VQA Natural Language Explanation via Self-Critical Learning, *Wei Suo, Mengyang Sun, Weisong Liu, Yiqi Gao, Peng Wang, Yanning Zhang, Qi Wu*
253. Teaching Structured Vision & Language Concepts to Vision & Language Models, *Sivan Doveh, Assaf Arbelle, Sivan Harary, Eli Schwartz, Roei Herzig, Raja Giryes, Rogerio Feris, Rameswar Panda, Shimon Ullman, Leonid Karlinsky*
254. FAME-ViL: Multi-Tasking Vision-Language Model for Heterogeneous Fashion Tasks, *Xiao Han, Xiatian Zhu, Licheng Yu, Li Zhang, Yi-Zhe Song, Tao Xiang*

255. RefCLIP: A Universal Teacher for Weakly Supervised Referring Expression Comprehension, *Lei Jin, Gen Luo, Yiyi Zhou, Xiaoshuai Sun, Guannan Jiang, Annan Shu, Rongrong Ji*
256. Uni-Perceiver v2: A Generalist Model for Large-Scale Vision and Vision-Language Tasks, *Hao Li, Jinguo Zhu, Xiaohu Jiang, Xizhou Zhu, Hongsheng Li, Chun Yuan, Xiaohua Wang, Yu Qiao, Xiaogang Wang, Wenhai Wang, Jifeng Dai*
257. Learning From Unique Perspectives: User-Aware Saliency Modeling, *Shi Chen, Nachiappan Valliappan, Shaolei Shen, Xinyu Ye, Kai Kohlhoff, Junfeng He*
258. CRAFT: Concept Recursive Activation FacTORIZATION for Explainability, *Thomas Fel, Agustin Picard, Louis Béthune, Thibaut Boissin, David Vigouroux, Julien Colin, Rémi Cadène, Thomas Serre*
259. Doubly Right Object Recognition: A Why Prompt for Visual Rationales, *Chengzhi Mao, Revant Teotia, Amrutha Sundar, Sachit Menon, Junfeng Yang, Xin Wang, Carl Vondrick*
260. Sketch2Saliency: Learning to Detect Salient Objects From Human Drawings, *Ayan Kumar Bhunia, Subhadeep Koley, Amandeep Kumar, Aneeshan Sain, Pinaki Nath Chowdhury, Tao Xiang, Yi-Zhe Song*
261. PIP-Net: Patch-Based Intuitive Prototypes for Interpretable Image Classification, *Meike Nauta, Jörg Schlötterer, Maurice van Keulen, Christin Seifert*
262. Photo Pre-Training, but for Sketch, *Ke Li, Kaiyue Pang, Yi-Zhe Song*
263. CLIP for All Things Zero-Shot Sketch-Based Image Retrieval, Fine-Grained or Not, *Aneeshan Sain, Ayan Kumar Bhunia, Pinaki Nath Chowdhury, Subhadeep Koley, Tao Xiang, Yi-Zhe Song*
264. iCLIP: Bridging Image Classification and Contrastive Language-Image Pre-Training for Visual Recognition, *Yixuan Wei, Yue Cao, Zheng Zhang, Houwen Peng, Zhuliang Yao, Zhenda Xie, Han Hu, Baining Guo*
265. Cross-Modal Implicit Relation Reasoning and Aligning for Text-to-Image Person Retrieval, *Ding Jiang, Mang Ye*
266. Multi-Modal Representation Learning With Text-Driven Soft Masks, *Jaeyoo Park, Bohyung Han*
267. Texts as Images in Prompt Tuning for Multi-Label Image Recognition, *Zixian Guo, Bowen Dong, Zhilong Ji, Jinfeng Bai, Yiwu Guo, Wangmeng Guo*
268. Reproducible Scaling Laws for Contrastive Language-Image Learning, *Mehdi Cherti, Romain Beaumont, Ross Wightman, Mitchell Wortsman, Gabriel Ilharco, Cade Gordon, Christoph Schuhmann, Ludwig Schmidt, Jenia Jitsev*
269. Multilateral Semantic Relations Modeling for Image Text Retrieval, *Zheng Wang, Zhenwei Gao, Kangshuai Guo, Yang Yang, Xiaoming Wang, Heng Tao Shen*
270. SmallCap: Lightweight Image Captioning Prompted With Retrieval Augmentation, *Rita Ramos, Bruno Martins, Desmond Elliott, Yova Kementchedjheva*
271. Probing Sentiment-Oriented Pre-Training Inspired by Human Sentiment Perception Mechanism, *Tinglei Feng, Jiaxuan Liu, Jufeng Yang*
272. Prefix Conditioning Unifies Language and Label Supervision, *Kuniaki Saito, Kihyuk Sohn, Xiang Zhang, Chun-Liang Li, Chen-Yu Lee, Kate Saenko, Tomas Pfister*
273. Crossing the Gap: Domain Generalization for Image Captioning, *Yuchen Ren, Zhendong Mao, Shancheng Fang, Yan Lu, Tong He, Hao Du, Yongdong Zhang, Wanli Ouyang*
274. A Bag-of-Prototypes Representation for Dataset-Level Applications, *Weijie Tu, Weijian Deng, Tom Gedeon, Liang Zheng*
275. CrowdCLIP: Unsupervised Crowd Counting via Vision-Language Model, *Dingkang Liang, Jiahao Xie, Zhikang Zou, Xiaoqing Ye, Wei Xu, Xiang Bai*
276. D²Former: Jointly Learning Hierarchical Detectors and Contextual Descriptors via Agent-Based Transformers, *Jianfeng He, Yuan Gao, Tianzhu Zhang, Zhe Zhang, Feng Wu*
277. Learning to Generate Language-Supervised and Open-Vocabulary Scene Graph Using Pre-Trained Visual-Semantic Space, *Yong Zhang, Yingwei Pan, Ting Yao, Rui Huang, Tao Mei, Chang-Wen Chen*
278. Relational Context Learning for Human-Object Interaction Detection, *Sanghyun Kim, Deunsol Jung, Minsu Cho*
279. Learning Open-Vocabulary Semantic Segmentation Models From Natural Language Supervision, *Jilan Xu, Junlin Hou, Yuejie Zhang, Rui Feng, Yi Wang, Yu Qiao, Weidi Xie*
280. Side Adapter Network for Open-Vocabulary Semantic Segmentation, *Mengde Xu, Zheng Zhang, Fangyun Wei, Han Hu, Xiang Bai*
281. Open-Vocabulary Panoptic Segmentation With Text-to-Image Diffusion Models, *Jiarui Xu, Sifei Liu, Arash Vahdat, Wonmin Byeon, Xiaolong Wang, Shalini De Mello*
282. IFSeg: Image-Free Semantic Segmentation via Vision-Language Model, *Sukmin Yun, Seong Hyeon Park, Paul Hongsuck Seo, Jinwoo Shin*
283. PartManip: Learning Cross-Category Generalizable Part Manipulation Policy From Point Cloud Observations, *Haoran Geng, Ziming Li, Yiran Geng, Jiayi Chen, Hao Dong, He Wang*
284. OneFormer: One Transformer to Rule Universal Image Segmentation, *Jitesh Jain, Jiachen Li, Mang Tik Chiu, Ali Hassani, Nikita Orlov, Humphrey Shi*
285. Delving Into Shape-Aware Zero-Shot Semantic Segmentation, *Xinyu Liu, Beiwen Tian, Zhen Wang, Rui Wang, Kehua Sheng, Bo Zhang, Hao Zhao, Guyue Zhou*
286. CoMFormer: Continual Learning in Semantic and Panoptic Segmentation, *Fabio Cermelli, Matthieu Cord, Arthur Douillard*
287. Learning to Segment Every Referring Object Point by Point, *Mengxue Qu, Yu Wu, Yunchao Wei, Wu Liu, Xiaodan Liang, Yao Zhao*
288. Unsupervised Continual Semantic Adaptation Through Neural Rendering, *Zhizheng Liu, Francesco Milano, Jonas Frey, Roland Siegwart, Hermann Blum, Cesar Cadena*
289. Mask DINO: Towards a Unified Transformer-Based Framework for Object Detection and Segmentation, *Feng Li, Hao Zhang, Huaizhe Xu, Shilong Liu, Lei Zhang, Lionel M. Ni, Heung-Yeung Shum*
290. Transformer Scale Gate for Semantic Segmentation, *Hengcan Shi, Munawar Hayat, Jianfei Cai*
291. Style Projected Clustering for Domain Generalized Semantic Segmentation, *Wei Huang, Chang Chen, Yong Li, Jiacheng Li, Cheng Li, Fenglong Song, Youliang Yan, Zhiwei Xiong*
292. Rethinking Few-Shot Medical Segmentation: A Vector Quantization View, *Shiqi Huang, Tingfa Xu, Ning Shen, Feng Mu, Jianan Li*
293. Continual Semantic Segmentation With Automatic Memory Sample Selection, *Lanyun Zhu, Tianrun Chen, Jianxiong Yin, Simon See, Jun Liu*

294. Token Contrast for Weakly-Supervised Semantic Segmentation, *Lixiang Ru, Heliang Zheng, Yibing Zhan, Bo Du*
295. Multi-Granularity Archaeological Dating of Chinese Bronze Dings Based on a Knowledge-Guided Relation Graph, *Rixin Zhou, Jiafu Wei, Qian Zhang, Ruihua Qi, Xi Yang, Chuntao Li*
296. Hunting Sparsity: Density-Guided Contrastive Learning for Semi-Supervised Semantic Segmentation, *Xiaoyang Wang, Bingfeng Zhang, Limin Yu, Jimin Xiao*
297. Cut and Learn for Unsupervised Object Detection and Instance Segmentation, *Xudong Wang, Rohit Girdhar, Stella X. Yu, Ishan Misra*
298. Extracting Class Activation Maps From Non-Discriminative Features As Well, *Zhaozheng Chen, Qianru Sun*
299. BoxTeacher: Exploring High-Quality Pseudo Labels for Weakly Supervised Instance Segmentation, *Tianheng Cheng, Xinggang Wang, Shaoyu Chen, Qian Zhang, Wenyu Liu*
300. Hierarchical Fine-Grained Image Forgery Detection and Localization, *Xiao Guo, Xiaohong Liu, Zhiyuan Ren, Steven Grosz, Iacopo Masi, Xiaoming Liu*
301. Towards Professional Level Crowd Annotation of Expert Domain Data, *Pei Wang, Nuno Vasconcelos*
302. Unsupervised Object Localization: Observing the Background to Discover Objects, *Oriane Siméoni, Chloé Sekkat, Gilles Puy, Antonín Vobecký, Éloi Zablocki, Patrick Pérez*
303. Semi-Supervised Learning Made Simple With Self-Supervised Clustering, *Enrico Fini, Pietro Astolfi, Karteek Alahari, Xavier Alameda-Pineda, Julien Mairal, Moin Nabi, Elisa Ricci*
304. Unbalanced Optimal Transport: A Unified Framework for Object Detection, *Henri De Plaen, Pierre-François De Plaen, Johan A. K. Suykens, Marc Proesmans, Tinne Tuytelaars, Luc Van Gool*
305. DiGeo: Discriminative Geometry-Aware Learning for Generalized Few-Shot Object Detection, *Jiawei Ma, Yulei Niu, Jincheng Xu, Shiyuan Huang, Guangxing Han, Shih-Fu Chang*
306. CLIP the Gap: A Single Domain Generalization Approach for Object Detection, *Vidit Vidit, Martin Engilberge, Mathieu Salzmann*
307. Unknown Sniffer for Object Detection: Don't Turn a Blind Eye to Unknown Objects, *Wenteng Liang, Feng Xue, Yihao Liu, Guofeng Zhong, Anlong Ming*
308. Consistent-Teacher: Towards Reducing Inconsistent Pseudo-
✧ Targets in Semi-Supervised Object Detection, *Xinjiang Wang, Xingyi Yang, Shilong Zhang, Yijiang Li, Litong Feng, Shijie Fang, Chengqi Lyu, Kai Chen, Wayne Zhang*
309. Optimal Proposal Learning for Deployable End-to-End Pedestrian Detection, *Xiaolin Song, Binghui Chen, Pengyu Li, Jun-Yan He, Biao Wang, Yifeng Geng, Xuansong Xie, Honggang Zhang*
310. AsyFOD: An Asymmetric Adaptation Paradigm for Few-Shot Domain Adaptive Object Detection, *Yipeng Gao, Kun-Yu Lin, Junkai Yan, Yaowei Wang, Wei-Shi Zheng*
311. Where Is My Spot? Few-Shot Image Generation via Latent Subspace Optimization, *Chenxi Zheng, Bangzhen Liu, Huaidong Zhang, Xuemiao Xu, Shengfeng He*
312. Uncertainty-Aware Optimal Transport for Semantically Coherent Out-of-Distribution Detection, *Fan Lu, Kai Zhu, Wei Zhai, Kecheng Zheng, Yang Cao*
313. MAESTER: Masked Autoencoder Guided Segmentation at Pixel
✧ Resolution for Accurate, Self-Supervised Subcellular Structure Recognition, *Ronald Xie, Kuan Pang, Gary D. Bader, Bo Wang*
314. Orthogonal Annotation Benefits Barely-Supervised Medical Image Segmentation, *Heng Cai, Shumeng Li, Lei Qi, Qian Yu, Yinghuan Shi, Yang Gao*
315. RepMode: Learning to Re-Parameterize Diverse Experts for
✧ Subcellular Structure Prediction, *Donghao Zhou, Chunbin Gu, Junde Xu, Furui Liu, Qiong Wang, Guangyong Chen, Pheng-Ann Heng*
316. Topology-Guided Multi-Class Cell Context Generation for Digital Pathology, *Shahira Abousamra, Rajarsi Gupta, Tahsin Kurc, Dimitris Samaras, Joel Saltz, Chao Chen*
317. Dynamic Graph Enhanced Contrastive Learning for Chest X-Ray Report Generation, *Mingjie Li, Bingqian Lin, Zicong Chen, Haokun Lin, Xiaodan Liang, Xiaojun Chang*
318. Benchmarking Self-Supervised Learning on Diverse Pathology Datasets, *Mingu Kang, Heon Song, Seonwook Park, Donggeun Yoo, Sérgio Pereira*
319. Multiple Instance Learning via Iterative Self-Paced Supervised Contrastive Learning, *Kangning Liu, Weicheng Zhu, Yiqiu Shen, Sheng Liu, Narges Razavian, Krzysztof J. Geras, Carlos Fernandez-Granda*
320. Learning Expressive Prompting With Residuals for Vision Transformers, *Rajshekhar Das, Yonatan Dukler, Avinash Ravichandran, Ashwin Swaminathan*
321. Detection of Out-of-Distribution Samples Using Binary Neuron Activation Patterns, *Bartłomiej Olber, Krystian Radlak, Adam Popowicz, Michal Szczepankiewicz, Krystian Chachula*
322. Decoupling MaxLogit for Out-of-Distribution Detection, *Zihan Zhang, Xiang Xiang*
323. Exploring Structured Semantic Prior for Multi Label Recognition With Incomplete Labels, *Zixuan Ding, Ao Wang, Hui Chen, Qiang Zhang, Pengzhang Liu, Yongjun Bao, Weipeng Yan, Jungong Han*
324. Bridging the Gap Between Model Explanations in Partially Annotated Multi-Label Classification, *Youngwook Kim, Jae Myung Kim, Jieun Jeong, Cordelia Schmid, Zeynep Akata, Jungwoo Lee*
325. DivClust: Controlling Diversity in Deep Clustering, *Ioannis Maniadis Metaxas, Georgios Tzimiropoulos, Ioannis Patras*
326. Deep Semi-Supervised Metric Learning With Mixed Label Propagation, *Furen Zhuang, Pierre Moulin*
327. Leveraging Inter-Rater Agreement for Classification in the Presence of Noisy Labels, *Maria Sofia Bucarelli, Lucas Cassano, Federico Siciliano, Amin Mantrach, Fabrizio Silvestri*
328. Modeling Inter-Class and Intra-Class Constraints in Novel Class Discovery, *Wenbin Li, Zhichen Fan, Jing Huo, Yang Gao*
329. Bootstrap Your Own Prior: Towards Distribution-Agnostic Novel Class Discovery, *Muli Yang, Liancheng Wang, Cheng Deng, Hanwang Zhang*
330. Towards Realistic Long-Tailed Semi-Supervised Learning: Consistency Is All You Need, *Tong Wei, Kai Gan*
331. PromptCAL: Contrastive Affinity Learning via Auxiliary Prompts for Generalized Novel Category Discovery, *Sheng Zhang, Salman Khan, Zhiqiang Shen, Muzammal Naseer, Guangyi Chen, Fahad Shahbaz Khan*
332. Probabilistic Knowledge Distillation of Face Ensembles, *Jianqing Xu, Shen Li, Ailin Deng, Miao Xiong, Jiaying Wu, Jiaxiang Wu, Shouhong Ding, Bryan Hooi*
333. Class-Conditional Sharpness-Aware Minimization for Deep Long-Tailed Recognition, *Zhipeng Zhou, Lanqing Li, Peilin Zhao, Pheng-Ann Heng, Wei Gong*

334. Promoting Semantic Connectivity: Dual Nearest Neighbors Contrastive Learning for Unsupervised Domain Generalization, *Yuchen Liu, Yaoming Wang, Yabo Chen, Wenrui Dai, Chenglin Li, Junni Zou, Hongkai Xiong*
335. Instance Relation Graph Guided Source-Free Domain Adaptive Object Detection, *Vibashan VS, Poojan Oza, Vishal M. Patel*
336. MOT: Masked Optimal Transport for Partial Domain Adaptation, *You-Wei Luo, Chuan-Xian Ren*
337. TOPLight: Lightweight Neural Networks With Task-Oriented Pretraining for Visible-Infrared Recognition, *Hao Yu, Xu Cheng, Wei Peng*
338. OSAN: A One-Stage Alignment Network to Unify Multimodal Alignment and Unsupervised Domain Adaptation, *Ye Liu, Lingfeng Qiao, Changchong Lu, Di Yin, Chen Lin, Haoyuan Peng, Bo Ren*
339. Patch-Mix Transformer for Unsupervised Domain Adaptation: A ♣ Game Perspective, *Jinjing Zhu, Haotian Bai, Lin Wang*
340. ARO-Net: Learning Implicit Fields From Anchored Radial Observations, *Yizhi Wang, Zeyu Huang, Ariel Shamir, Hui Huang, Hao Zhang, Ruizhen Hu*
341. A Probabilistic Framework for Lifelong Test-Time Adaptation, *Dhanajit Brahma, Piyush Rai*
342. Distribution Shift Inversion for Out-of-Distribution Prediction, *Runpeng Yu, Songhua Liu, Xingyi Yang, Xinchao Wang*
343. Learning Joint Latent Space EBM Prior Model for Multi-Layer Generator, *Jiali Cui, Ying Nian Wu, Tian Han*
344. A Data-Based Perspective on Transfer Learning, *Saachi Jain, Hadi Salman, Alaa Khaddaj, Eric Wong, Sung Min Park, Aleksander Mqdry*
345. A Meta-Learning Approach to Predicting Performance and Data Requirements, *Achin Jain, Gurumurthy Swaminathan, Paolo Favaro, Hao Yang, Avinash Ravichandran, Hrayr Harutyunyan, Alessandro Achille, Onkar Dabeer, Bernt Schiele, Ashwin Swaminathan, Stefano Soatto*
346. Guided Recommendation for Model Fine-Tuning, *Hao Li, Charless Fowlkes, Hao Yang, Onkar Dabeer, Zhuowen Tu, Stefano Soatto*
347. EMT-NAS: Transferring Architectural Knowledge Between Tasks From Different Datasets, *Peng Liao, Yaochu Jin, Wenli Du*
348. AttrICLIP: A Non-Incremental Learner for Incremental Knowledge Learning, *Runqi Wang, Xiaoyue Duan, Guoliang Kang, Jianzhuang Liu, Shaohui Lin, Songcen Xu, Jinhu Lü, Baochang Zhang*
349. Batch Model Consolidation: A Multi-Task Model Consolidation Framework, *Iordanis Fostropoulos, Jiaye Zhu, Laurent Itti*
350. SmartAssign: Learning a Smart Knowledge Assignment Strategy for Deraining and Desnowing, *Yinglong Wang, Chao Ma, Jianzhuang Liu*
351. TinyMIM: An Empirical Study of Distilling MIM Pre-Trained Models, *Sucheng Ren, Fangyun Wei, Zheng Zhang, Han Hu*
352. Computationally Budgeted Continual Learning: What Does Matter? *Ameya Prabhu, Hasan Abed Al Kader Hammoud, Puneet K. Dokania, Philip H.S. Torr, Ser-Nam Lim, Bernard Ghanem, Adel Bibi*
353. GradMA: A Gradient-Memory-Based Accelerated Federated ♣ Learning With Alleviated Catastrophic Forgetting, *Kangyang Luo, Xiang Li, Yunshi Lan, Ming Gao*
354. Rethinking Gradient Projection Continual Learning: Stability / Plasticity Feature Space Decoupling, *Zhen Zhao, Zhizhong Zhang, Xin Tan, Jun Liu, Yanyun Qu, Yuan Xie, Lizhuang Ma*
355. Neuro-Modulated Hebbian Learning for Fully Test-Time Adaptation, *Yushun Tang, Ce Zhang, Heng Xu, Shuoshuo Chen, Jie Cheng, Luziwei Leng, Qinghai Guo, Zhihai He*
356. Generalizing Dataset Distillation via Deep Generative Prior, *George Cazenavette, Tongzhou Wang, Antonio Torralba, Alexei A. Efros, Jun-Yan Zhu*
357. Minimizing the Accumulated Trajectory Error to Improve Dataset Distillation, *Jiawei Du, Yidi Jiang, Vincent Y. F. Tan, Joey Tianyi Zhou, Haizhou Li*
358. Slimmable Dataset Condensation, *Songhua Liu, Jingwen Ye, ♣ Runpeng Yu, Xinchao Wang*
359. Sharpness-Aware Gradient Matching for Domain Generalization, *Pengfei Wang, Zhaoxiang Zhang, Zhen Lei, Lei Zhang*
360. Dynamic Neural Network for Multi-Task Learning Searching Across Diverse Network Topologies, *Wonhyeok Choi, Sunghoon Im*
361. SplineCam: Exact Visualization and Characterization of Deep ♣ Network Geometry and Decision Boundaries, *Ahmed Imtiaz Humayun, Randall Balestriero, Guha Balakrishnan, Richard G. Baraniuk*
362. VNE: An Effective Method for Improving Deep Representation by Manipulating Eigenvalue Distribution, *Jaeill Kim, Suhyun Kang, Duhun Hwang, Jungwook Shin, Wonjong Rhee*
363. Efficient On-Device Training via Gradient Filtering, *Yuedong Yang, Guihong Li, Radu Marculescu*
364. Are Data-Driven Explanations Robust Against Out-of-Distribution Data? *Tang Li, Fengchun Qiao, Mengmeng Ma, Xi Peng*
365. BiasAdv: Bias-Adversarial Augmentation for Model Debiasing, *Jongin Lim, Youngdong Kim, Byungjai Kim, Chanho Ahn, Jinwoo Shin, Eunho Yang, Seungju Han*
366. Q-DETR: An Efficient Low-Bit Quantized Detection ♣ Transformer, *Sheng Xu, Yanjing Li, Mingbao Lin, Peng Gao, Guodong Guo, Jinhu Lü, Baochang Zhang*
367. NIPQ: Noise Proxy-Based Integrated Pseudo-Quantization, *Juncheol Shin, Junhyuk So, Sein Park, Seungyeop Kang, Sungjoo Yoo, Eunhyeok Park*
368. CUDA: Convolution-Based Unlearnable Datasets, *Vinu Sankar Sadasivan, Mahdi Soltanolkotabi, Soheil Feizi*
369. KD-DLGAN: Data Limited Image Generation via Knowledge Distillation, *Kaiwen Cui, Yingchen Yu, Fangneng Zhan, Shengcai Liao, Shijian Lu, Eric P. Xing*
370. Spider GAN: Leveraging Friendly Neighbors to Accelerate GAN Training, *Siddharth Asokan, Chandra Sekhar Seelamantula*
371. Efficient Verification of Neural Networks Against LVM-Based Specifications, *Harleen Hanspal, Alessio Lomuscio*
372. Bi-Directional Feature Fusion Generative Adversarial Network for Ultra-High Resolution Pathological Image Virtual Re-Staining, *Kexin Sun, Zhineng Chen, Gongwei Wang, Jun Liu, Xiongjun Ye, Yu-Gang Jiang*
373. DeSTSeg: Segmentation Guided Denoising Student-Teacher for Anomaly Detection, *Xuan Zhang, Shiyu Li, Xi Li, Ping Huang, Jiulong Shan, Ting Chen*
374. OmniAL: A Unified CNN Framework for Unsupervised Anomaly Localization, *Ying Zhao*

- Notes:**

[illegible][illegible]

1230–1400 Poster Switch/Setup (West Exhibit Hall)

1400–1600 Demos (West Exhibit Hall Demo Area)

- Same as morning demos (see page 3).

1400–1500 Plenary-Tue-PM (East Exhibit Halls A–B)

Panel: History and Future of Artificial Intelligence and Computer Vision

Moderator: Svetlana Lazebnik (*Univ. of Illinois at Urbana-Champaign*)

Participants:

- Chelsea Finn (*Stanford*)
- Dan Huttenlocher (*MIT*)
- Linda Shapiro (*Univ. of Washington*)
- Jamie Shotton (*Wayve*)

1500–1600 Oral-Award-Tue (East Exhibit Halls A–B)

Award candidate paper presentations

Format: 10 min. presentation (including questions)

1. **[1500]** Ego-Body Pose Estimation via Ego-Head Pose Estimation, *Jiaman Li, Karen Liu, Jiajun Wu*
2. **[1510]** 3D Registration With Maximal Cliques, *Xiyu Zhang, Jiaqi Yang, Shikun Zhang, Yanning Zhang*
3. **[1520]** OmniObject3D: Large-Vocabulary 3D Object Dataset for Realistic Perception, Reconstruction and Generation, *Tong Wu, Jiarui Zhang, Xiao Fu, Yuxin Wang, Jiawei Ren, Liang Pan, Wayne Wu, Lei Yang, Jiaqi Wang, Chen Qian, Dahua Lin, Ziwei Liu*
4. **[1530]** MobileNeRF: Exploiting the Polygon Rasterization Pipeline for Efficient Neural Field Rendering on Mobile Architectures, *Zhiqin Chen, Thomas Funkhouser, Peter Hedman, Andrea Tagliasacchi*
5. **[1540]** DynBaR: Neural Dynamic Image-Based Rendering, *Zhengqi Li, Qianqian Wang, Forrester Cole, Richard Tucker, Noah Snavely*
6. **[1550]** Planning-Oriented Autonomous Driving, *Yihan Hu, Jiazhi Yang, Li Chen, Keyu Li, Chonghao Sima, Xizhou Zhu, Siqi Chai, Senyao Du, Tianwei Lin, Wenhai Wang, Lewei Lu, Xiaosong Jia, Qiang Liu, Jifeng Dai, Yu Qiao, Hongyang Li*

1600–1700 Afternoon Break (West Exhibit Hall)

1630–1830 Poster-Tue-PM (West Exhibit Hall)

✧ - Highlight paper (check it out)

☞ - Award candidate paper (see award sessions)

1. Passive Micron-Scale Time-of-Flight With Sunlight
✧ Interferometry, *Alankar Kotwal, Anat Levin, Ioannis Gkioulekas*
2. F²-NeRF: Fast Neural Radiance Field Training With Free Camera Trajectories, *Peng Wang, Yuan Liu, Zhaoxi Chen, Lingjie Liu, Ziwei Liu, Taku Komura, Christian Theobalt, Wenping Wang*
3. NoPe-NeRF: Optimising Neural Radiance Field With No Pose
✧ Prior, *Wenjing Bian, Zirui Wang, Kejie Li, Jia-Wang Bian, Victor Adrian Prisacariu*
4. BAD-NeRF: Bundle Adjusted Deblur Neural Radiance Fields, *Peng Wang, Lingzhe Zhao, Ruijie Ma, Peidong Liu*
5. DiffusioNeRF: Regularizing Neural Radiance Fields With Denoising Diffusion Models, *Jamie Wynn, Daniyar Turmukhambetov*

6. SPARF: Neural Radiance Fields From Sparse and Noisy Poses,
✧ ✧ *Prune Truong, Marie-Julie Rakotosaona, Fabian Manhardt, Federico Tombari*
7. Interactive Segmentation of Radiance Fields, *Rahul Goel, Dhawal Sirikonda, Saurabh Saini, P. J. Narayanan*
8. Temporal Interpolation Is All You Need for Dynamic Neural Radiance Fields, *Sungheon Park, Minjung Son, Seokhwan Jang, Young Chun Ahn, Ji-Yeon Kim, Nahyup Kang*
9. Compressing Volumetric Radiance Fields to 1 MB, *Lingzhi Li, Zhen Shen, Zhongshu Wang, Li Shen, Liefeng Bo*
10. Multiscale Tensor Decomposition and Rendering Equation Encoding for View Synthesis, *Kang Han, Wei Xiang*
11. Ref-NPR: Reference-Based Non-Photorealistic Radiance Fields for Controllable Scene Stylization, *Yuechen Zhang, Zexin He, Jinbo Xing, Xufeng Yao, Jiaya Jia*
12. Representing Volumetric Videos As Dynamic MLP Maps, *Sida Peng, Yunzhi Yan, Qing Shuai, Hujun Bao, Xiaowei Zhou*
13. Fast Monocular Scene Reconstruction With Global-Sparse Local-Dense Grids, *Wei Dong, Christopher Choy, Charles Loop, Or Litany, Yuke Zhu, Anima Anandkumar*
14. DynBaR: Neural Dynamic Image-Based Rendering, *Zhengqi Li, Qianqian Wang, Forrester Cole, Richard Tucker, Noah Snavely*
15. Plateau-Reduced Differentiable Path Tracing, *Michael Fischer, Tobias Ritschel*
16. NeFII: Inverse Rendering for Reflectance Decomposition With Near-Field Indirect Illumination, *Haoqian Wu, Zhipeng Hu, Lincheng Li, Yongqiang Zhang, Changjie Fan, Xin Yu*
17. WildLight: In-the-Wild Inverse Rendering With a Flashlight, *Ziang Cheng, Junxuan Li, Hongdong Li*
18. Relightable Neural Human Assets From Multi-View Gradient Illuminations, *Taotao Zhou, Kai He, Di Wu, Teng Xu, Qixuan Zhang, Kuixiang Shao, Wenzheng Chen, Lan Xu, Jingyi Yu*
19. DiffRF: Rendering-Guided 3D Radiance Field Diffusion, *Norman Müller, Yawar Siddiqui, Lorenzo Porzi, Samuel Rota Bulò, Peter Kotschieder, Matthias Nießner*
20. Analyzing Physical Impacts Using Transient Surface Wave Imaging, *Tianyuan Zhang, Mark Sheinin, Dorian Chan, Mark Rau, Matthew O'Toole, Srinivasa G. Narasimhan*
21. Neural Kaleidoscopic Space Sculpting, *Byeongjoo Ahn, Michael De Zeeuw, Ioannis Gkioulekas, Aswin C. Sankaranarayanan*
22. Towards Unbiased Volume Rendering of Neural Implicit Surfaces With Geometry Priors, *Yongqiang Zhang, Zhipeng Hu, Haoqian Wu, Minda Zhao, Lincheng Li, Zhengxia Zou, Changjie Fan*
23. Neural Kernel Surface Reconstruction, *Jiahui Huang, Zan Gojcic, Matan Atzmon, Or Litany, Sanja Fidler, Francis Williams*
24. MM-3DScene: 3D Scene Understanding by Customizing Masked Modeling With Informative-Preserved Reconstruction and Self-Distilled Consistency, *Mingye Xu, Mutian Xu, Tong He, Wanli Ouyang, Yali Wang, Xiaoguang Han, Yu Qiao*
25. Shape, Pose, and Appearance From a Single Image via Bootstrapped Radiance Field Inversion, *Dario Pavllo, David Joseph Tan, Marie-Julie Rakotosaona, Federico Tombari*
26. DisCoScene: Spatially Disentangled Generative Radiance Fields for Controllable 3D-Aware Scene Synthesis, *Yinghao Xu, Menglei Chai, Zifan Shi, Sida Peng, Ivan Skorokhodov, Aliaksandr Siarohin, Ceyuan Yang, Yujun Shen, Hsin-Ying Lee, Bolei Zhou, Sergey Tulyakov*

27. Heat Diffusion Based Multi-Scale and Geometric Structure-Aware Transformer for Mesh Segmentation, *Chi-Chong Wong*
28. Learning Detailed Radiance Manifolds for High-Fidelity and 3D-Consistent Portrait Synthesis From Monocular Image, *Yu Deng, Baoyuan Wang, Heung-Yeung Shum*
29. 3D-Aware Conditional Image Synthesis, *Kangle Deng, Gengshan Yang, Deva Ramanan, Jun-Yan Zhu*
30. VIVE3D: Viewpoint-Independent Video Editing Using 3D-Aware GANs, *Anna Frühstück, Nikolaos Sarafianos, Yuanlu Xu, Peter Wonka, Tony Tung*
31. SDFusion: Multimodal 3D Shape Completion, Reconstruction, and Generation, *Yen-Chi Cheng, Hsin-Ying Lee, Sergey Tulyakov, Alexander G. Schwing, Liang-Yan Gui*
32. Generating Part-Aware Editable 3D Shapes Without 3D Supervision, *Konstantinos Tertikas, Despoina Paschalidou, Boxiao Pan, Jeong Joon Park, Mikaela Angelina Uy, Ioannis Emiris, Yannis Avrithis, Leonidas Guibas*
33. NeuralLift-360: Lifting an In-the-Wild 2D Photo to a 3D Object
✧ With 360° Views, *Dejia Xu, Yifan Jiang, Peihao Wang, Zhiwen Fan, Yi Wang, Zhangyang Wang*
34. Implicit Identity Driven Deepfake Face Swapping Detection, *Baojin Huang, Zhongyuan Wang, Jifan Yang, Jiaxin Ai, Qin Zou, Qian Wang, Dengpan Ye*
35. Canonical Fields: Self-Supervised Learning of Pose-
✧ Canonicalized Neural Fields, *Rohith Agaram, Shaurya Dewan, Rahul Sajnani, Adrien Poulenard, Madhava Krishna, Srinath Sridhar*
36. Improving Fairness in Facial Albedo Estimation via Visual-
✧ Textual Cues, *Xingyu Ren, Jiankang Deng, Chao Ma, Yichao Yan, Xiaokang Yang*
37. High-Fidelity 3D Face Generation From Natural Language Descriptions, *Menghua Wu, Hao Zhu, Linjia Huang, Yiyu Zhuang, Yuanxun Lu, Xun Cao*
38. DSFNet: Dual Space Fusion Network for Occlusion-Robust 3D Dense Face Alignment, *Heyuan Li, Bo Wang, Yu Cheng, Mohan Kankanhalli, Robby T. Tan*
39. High-Fidelity Facial Avatar Reconstruction From Monocular Video With Generative Priors, *Yunpeng Bai, Yanbo Fan, Xuan Wang, Yong Zhang, Jingxiang Sun, Chun Yuan, Ying Shan*
40. 3DAvatarGAN: Bridging Domains for Personalized Editable Avatars, *Rameen Abdal, Hsin-Ying Lee, Peihao Zhu, Menglei Chai, Aliaksandr Siarohin, Peter Wonka, Sergey Tulyakov*
41. RODIN: A Generative Model for Sculpting 3D Digital Avatars
✧ Using Diffusion, *Tengfei Wang, Bo Zhang, Ting Zhang, Shuyang Gu, Jianmin Bao, Tadas Baltrusaitis, Jingjing Shen, Dong Chen, Fang Wen, Qifeng Chen, Baining Guo*
42. Instant Volumetric Head Avatars, *Wojciech Zielonka, Timo Bolkart, Justus Thies*
43. Synthesizing Photorealistic Virtual Humans Through Cross-Modal Disentanglement, *Siddharth Ravichandran, Ondřej Texler, Dimitar Dinev, Hyun Jae Kang*
44. 3D Cinemagraphy From a Single Image, *Xingyi Li, Zhiguo Cao, Huiqiang Sun, Jianming Zhang, Ke Xian, Guosheng Lin*
45. TryOnDiffusion: A Tale of Two UNets, *Luyang Zhu, Dawei Yang, Tyler Zhu, Fitsum Reda, William Chan, Chitwan Saharia, Mohammad Norouzi, Ira Kemelmacher-Shlizerman*
46. Diverse 3D Hand Gesture Prediction From Body Dynamics by Bilateral Hand Disentanglement, *Xingqun Qi, Chen Liu, Muyi Sun, Lincheng Li, Changjie Fan, Xin Yu*
47. Normal-Guided Garment UV Prediction for Human Re-
✧ Texturing, *Yasamin Jafarian, Tuanfeng Y. Wang, Duygu Ceylan, Jimei Yang, Nathan Carr, Yi Zhou, Hyun Soo Park*
48. REC-MV: REconstructing 3D Dynamic Cloth From Monocular Videos, *Lingteng Qiu, Guanying Chen, Jiapeng Zhou, Mutian Xu, Junle Wang, Xiaoguang Han*
49. SeSDF: Self-Evolved Signed Distance Field for Implicit 3D Clothed Human Reconstruction, *Yukang Cao, Kai Han, Kwan-Yee K. Wong*
50. Unsupervised Volumetric Animation, *Aliaksandr Siarohin, Willi Menapace, Ivan Skorokhodov, Kyle Olszewski, Jian Ren, Hsin-Ying Lee, Menglei Chai, Sergey Tulyakov*
51. Handy: Towards a High Fidelity 3D Hand Shape and Appearance Model, *Rolandos Alexandros Potamias, Stylianos Ploumpis, Stylianos Moschoglou, Vasileios Triantafyllou, Stefanos Zafeiriou*
52. Fantastic Breaks: A Dataset of Paired 3D Scans of Real-World Broken Objects and Their Complete Counterparts, *Nikolas Lamb, Cameron Palmer, Benjamin Molloy, Sean Banerjee, Natasha Kholgade Banerjee*
53. Distilling Neural Fields for Real-Time Articulated Shape Reconstruction, *Jeff Tan, Gengshan Yang, Deva Ramanan*
54. GANmouflage: 3D Object Nondetection With Texture Fields, *Rui Guo, Jasmine Collins, Oscar de Lima, Andrew Owens*
55. 3D Human Pose Estimation via Intuitive Physics, *Shashank Tripathi, Lea Müller, Chun-Hao P. Huang, Omid Taheri, Michael J. Black, Dimitrios Tzionas*
56. Object Pop-Up: Can We Infer 3D Objects and Their Poses From Human Interactions Alone? *Ilya A. Petrov, Riccardo Marin, Julian Chibane, Gerard Pons-Moll*
57. UniDexGrasp: Universal Robotic Dexterous Grasping via Learning Diverse Proposal Generation and Goal-Conditioned Policy, *Yinzheng Xu, Weikang Wan, Jialiang Zhang, Haoran Liu, Zikang Shan, Hao Shen, Ruicheng Wang, Haoran Geng, Yijia Weng, Jiayi Chen, Tengyu Liu, Li Yi, He Wang*
58. Constrained Evolutionary Diffusion Filter for Monocular Endoscope Tracking, *Xiongbiao Luo*
59. Visibility Aware Human-Object Interaction Tracking From Single RGB Camera, *Xianghui Xie, Bharat Lal Bhatnagar, Gerard Pons-Moll*
60. Transformer-Based Unified Recognition of Two Hands Manipulating Objects, *Hoseong Cho, Chanwoo Kim, Jihyeon Kim, Seongyeong Lee, Elkhani Ismayilzada, Seungryul Baek*
61. HuManiFlow: Ancestor-Conditioned Normalising Flows on SO(3) Manifolds for Human Pose and Shape Distribution Estimation, *Akash Sengupta, Ignas Budvytis, Roberto Cipolla*
62. 3D Human Pose Estimation With Spatio-Temporal Criss-Cross Attention, *Zhenhua Tang, Zhaofan Qiu, Yanbin Hao, Richang Hong, Ting Yao*
63. GFpose: Learning 3D Human Pose Prior With Gradient Fields, *Hai Ci, Mingdong Wu, Wentao Zhu, Xiaoxuan Ma, Hao Dong, Fangwei Zhong, Yizhou Wang*
64. JRDB-Pose: A Large-Scale Dataset for Multi-Person Pose Estimation and Tracking, *Edward Vendrow, Duy Tho Le, Jianfei Cai, Hamid Rezatofighi*
65. Analyzing and Diagnosing Pose Estimation With Attributions, *Qiyuan He, Linlin Yang, Kerui Gu, Qiuxia Lin, Angela Yao*
66. Shape-Constraint Recurrent Flow for 6D Object Pose Estimation, *Yang Hai, Rui Song, Jiaojiao Li, Yinlin Hu*

67. TexPose: Neural Texture Learning for Self-Supervised 6D Object Pose Estimation, *Hanzhi Chen, Fabian Manhardt, Nassir Navab, Benjamin Busam*
68. Hi-LASSIE: High-Fidelity Articulated Shape and Skeleton Discovery From Sparse Image Ensemble, *Chun-Han Yao, Wei-Chih Hung, Yuanzhen Li, Michael Rubinstein, Ming-Hsuan Yang, Varun Jampani*
69. Revisiting Rolling Shutter Bundle Adjustment: Toward Accurate and Fast Solution, *Bangyan Liao, Delin Qu, Yifei Xue, Huiqing Zhang, Yizhen Lao*
70. Revisiting the P3P Problem, *Yaqing Ding, Jian Yang, Viktor Larsson, Carl Olsson, Kalle Åström*
71. Common Pets in 3D: Dynamic New-View Synthesis of Real-Life ✧ Deformable Categories, *Samarth Sinha, Roman Shapovalov, Jeremy Reizenstein, Ignacio Rocco, Natalia Neverova, Andrea Vedaldi, David Novotny*
72. MobileBrick: Building LEGO for 3D Reconstruction on Mobile Devices, *Kejie Li, Jia-Wang Bian, Robert Castle, Philip H.S. Torr, Victor Adrian Prisacariu*
73. EFEM: Equivariant Neural Field Expectation Maximization for 3D Object Segmentation Without Scene Supervision, *Jiahui Lei, Congyue Deng, Karl Schmeckpeper, Leonidas Guibas, Kostas Daniilidis*
74. GINA-3D: Learning to Generate Implicit Neural Assets in the Wild, *Bokui Shen, Xinchun Yan, Charles R. Qi, Mahyar Najibi, Boyang Deng, Leonidas Guibas, Yin Zhou, Dragomir Anguelov*
75. Habitat-Matterport 3D Semantics Dataset, *Karmesh Yadav, ✧ Ram Ramrakhya, Santhosh Kumar Ramakrishnan, Theo Gervet, John Turner, Aaron Gokaslan, Noah Maestre, Angel Xuan Chang, Dhruv Batra, Manolis Savva, Alexander William Clegg, Devendra Singh Chaplot*
76. BUOL: A Bottom-Up Framework With Occupancy-Aware Lifting for Panoptic 3D Scene Reconstruction From a Single Image, *Tao Chu, Pan Zhang, Qiong Liu, Jiaqi Wang*
77. Panoptic Compositional Feature Field for Editable Scene Rendering With Network-Inferred Labels via Metric Learning, *Xinhua Cheng, Yanmin Wu, Mengxi Jia, Qian Wang, Jian Zhang*
78. A Light Touch Approach to Teaching Transformers Multi-View Geometry, *Yash Bhalgat, João F. Henriques, Andrew Zisserman*
79. Learning to Render Novel Views From Wide-Baseline Stereo Pairs, *Yilun Du, Cameron Smith, Ayush Tewari, Vincent Sitzmann*
80. Spring: A High-Resolution High-Detail Dataset and Benchmark for Scene Flow, Optical Flow and Stereo, *Lukas Mehl, Jenny Schmalfuss, Azin Jahedi, Yaroslava Nalivayko, Andrés Bruhn*
81. EventNeRF: Neural Radiance Fields From a Single Colour Event Camera, *Viktor Rudnev, Mohamed Elgharib, Christian Theobalt, Vladislav Golyanik*
82. LightedDepth: Video Depth Estimation in Light of Limited Inference View Angles, *Shengjie Zhu, Xiaoming Liu*
83. Generating Aligned Pseudo-Supervision From Non-Aligned Data for Image Restoration in Under-Display Camera, *Ruicheng Feng, Chongyi Li, Huaijin Chen, Shuai Li, Jinwei Gu, Chen Change Loy*
84. Spatio-Focal Bidirectional Disparity Estimation From a Dual-Pixel Image, *Donggun Kim, Hyeonjoong Jang, Inchul Kim, Min H. Kim*
85. Trap Attention: Monocular Depth Estimation With Manual Traps, *Chao Ning, Hongping Gan*
86. Accelerated Coordinate Encoding: Learning to Relocalize in ✧ Minutes Using RGB and Poses, *Eric Brachmann, Tommaso Cavallari, Victor Adrian Prisacariu*
87. Energy-Efficient Adaptive 3D Sensing, *Brevin Tilmon, Zhanghao Sun, Sanjeev J. Koppal, Yicheng Wu, Georgios Evangelidis, Ramzi Zahreddine, Gurunandan Krishnan, Sizhuo Ma, Jian Wang*
88. Incremental 3D Semantic Scene Graph Prediction From RGB Sequences, *Shun-Cheng Wu, Keisuke Tateno, Nassir Navab, Federico Tombari*
89. Consistent Direct Time-of-Flight Video Depth Super-Resolution, *Zhanghao Sun, Wei Ye, Jinhui Xiong, Gyeongmin Choe, Jialiang Wang, Shuochen Su, Rakesh Ranjan*
90. Learning to Zoom and Unzoom, *Chittesh Thavamani, Mengtian Li, Francesco Ferroni, Deva Ramanan*
91. FrustumFormer: Adaptive Instance-Aware Resampling for Multi-View 3D Detection, *Yuqi Wang, Yuntao Chen, Zhaoxiang Zhang*
92. 3D Video Object Detection With Learnable Object-Centric Global Optimization, *Jiawei He, Yuntao Chen, Naiyan Wang, Zhaoxiang Zhang*
93. UniDistill: A Universal Cross-Modality Knowledge Distillation ✧ Framework for 3D Object Detection in Bird's-Eye View, *Shengchao Zhou, Weizhou Liu, Chen Hu, Shuchang Zhou, Chao Ma*
94. ARKitTrack: A New Diverse Dataset for Tracking Using Mobile RGB-D Data, *Haojie Zhao, Junsong Chen, Lijun Wang, Huchuan Lu*
95. Deep Dive Into Gradients: Better Optimization for 3D Object Detection With Gradient-Corrected IoU Supervision, *Qi Ming, Lingjuan Miao, Zhe Ma, Lin Zhao, Zhiqiang Zhou, Xuhui Huang, Yuanpei Chen, Yufei Guo*
96. SlowLiDAR: Increasing the Latency of LiDAR-Based Detection Using Adversarial Examples, *Han Liu, Yuhao Wu, Zhiyuan Yu, Yevgeniy Vorobeychik, Ning Zhang*
97. Normalizing Flow Based Feature Synthesis for Outlier-Aware ✧ Object Detection, *Nishant Kumar, Siniša Šegvić, Abouzar Eslami, Stefan Gumhold*
98. OcTr: Octree-Based Transformer for 3D Object Detection, *Chao Zhou, Yanan Zhang, Jiaxin Chen, Di Huang*
99. HypLiLoc: Towards Effective LiDAR Pose Regression With Hyperbolic Fusion, *Sijie Wang, Qiyu Kang, Rui She, Wei Wang, Kai Zhao, Yang Song, Wee Peng Tay*
100. LiDAR2Map: In Defense of LiDAR-Based Semantic Map Construction Using Online Camera Distillation, *Song Wang, Wentong Li, Wenyu Liu, Xiaolu Liu, Jianke Zhu*
101. MSF: Motion-Guided Sequential Fusion for Efficient 3D Object Detection From Point Cloud Sequences, *Chenhang He, Ruihuang Li, Yabin Zhang, Shuai Li, Lei Zhang*
102. SFD2: Semantic-Guided Feature Detection and Description, *Fei Xue, Ignas Budvytis, Roberto Cipolla*
103. Temporal Consistent 3D LiDAR Representation Learning for Semantic Perception in Autonomous Driving, *Lucas Nunes, Louis Wiesmann, Rodrigo Marcuzzi, Xieyuanli Chen, Jens Behley, Cyrill Stachniss*
104. Unsupervised 3D Point Cloud Representation Learning by Triangle Constrained Contrast for Autonomous Driving, *Bo Pang, Hongchi Xia, Cewu Lu*

105. RangeViT: Towards Vision Transformers for 3D Semantic Segmentation in Autonomous Driving, *Angelika Ando, Spyros Gidaris, Andrei Bursuc, Gilles Puy, Alexandre Boulch, Renaud Marlet*
106. Spatiotemporal Self-Supervised Learning for Point Clouds in the Wild, *Yanhao Wu, Tong Zhang, Wei Ke, Sabine Süsstrunk, Mathieu Salzmann*
107. Change-Aware Sampling and Contrastive Learning for Satellite Images, *Utkarsh Mall, Bharath Hariharan, Kavita Bala*
108. Self-Supervised 3D Scene Flow Estimation Guided by Superpoints, *Yaqi Shen, Le Hui, Jin Xie, Jian Yang*
109. SCOOP: Self-Supervised Correspondence and Optimization-Based Scene Flow, *Itai Lang, Dror Aiger, Forrester Cole, Shai Avidan, Michael Rubinstein*
110. PiMAE: Point Cloud and Image Interactive Masked Autoencoders for 3D Object Detection, *Anthony Chen, Kevin Zhang, Renrui Zhang, Zihan Wang, Yuheng Lu, Yandong Guo, Shanghang Zhang*
111. CP³: Channel Pruning Plug-In for Point-Based Networks, *Yaomin Huang, Ning Liu, Zhengping Che, Zhiyuan Xu, Chaomin Shen, Yaxin Peng, Guixu Zhang, Xinmei Liu, Feifei Feng, Jian Tang*
112. Binarizing Sparse Convolutional Networks for Efficient Point Cloud Analysis, *Xiuwei Xu, Ziwei Wang, Jie Zhou, Jiwen Lu*
113. Hyperspherical Embedding for Point Cloud Completion, *Junming Zhang, Haomeng Zhang, Ram Vasudevan, Matthew Johnson-Roberson*
114. Attention-Based Point Cloud Edge Sampling, *Chengzhi Wu, Junwei Zheng, Julius Pfommer, Jürgen Beyer*
115. Starting From Non-Parametric Networks for 3D Point Cloud Analysis, *Renrui Zhang, Lihui Wang, Yali Wang, Peng Gao, Hongsheng Li, Jianbo Shi*
116. Grad-PU: Arbitrary-Scale Point Cloud Upsampling via Gradient Descent With Learned Distance Functions, *Yun He, Danhang Tang, Yinda Zhang, Xiangyang Xue, Yanwei Fu*
117. SE-ORNet: Self-Ensembling Orientation-Aware Network for Unsupervised Point Cloud Shape Correspondence, *Jiacheng Deng, Chuxin Wang, Jiahao Lu, Jianfeng He, Tianzhu Zhang, Jiyang Yu, Zhe Zhang*
118. Robust 3D Shape Classification via Non-Local Graph Attention Network, *Shengwei Qin, Zhong Li, Ligang Liu*
119. Rotation-Invariant Transformer for Point Cloud Matching, *Hao Yu, Zheng Qin, Ji Hou, Mahdi Saleh, Dongsheng Li, Benjamin Busam, Slobodan Ilic*
120. Deep Graph-Based Spatial Consistency for Robust Non-Rigid Point Cloud Registration, *Zheng Qin, Hao Yu, Changjian Wang, Yuxing Peng, Kai Xu*
121. Efficient RGB-T Tracking via Cross-Modality Distillation, *Tianlu Zhang, Hongyuan Guo, Qiang Jiao, Qiang Zhang, Jungong Han*
122. Finding Geometric Models by Clustering in the Consensus Space, *Daniel Barath, Denys Rozumnyi, Ivan Eichhardt, Levente Hajder, Jiri Matas*
123. Adaptive Assignment for Geometry Aware Local Feature Matching, *Dihe Huang, Ying Chen, Yong Liu, Jianlin Liu, Shang Xu, Wenlong Wu, Yikang Ding, Fan Tang, Chengjie Wang*
124. Masked Representation Learning for Domain Generalized Stereo Matching, *Zhibo Rao, Bangshu Xiong, Mingyi He, Yuchao Dai, Renjie He, Zhelun Shen, Xing Li*
125. Learning Optical Expansion From Scale Matching, *Han Ling, Yinghui Sun, Quansen Sun, Zhenwen Ren*
126. AnyFlow: Arbitrary Scale Optical Flow With Implicit Neural Representation, *Hyunyoung Jung, Zhuo Hui, Lei Luo, Haitao Yang, Feng Liu, Sungjoo Yoo, Rakesh Ranjan, Denis Demandolx*
127. HouseDiffusion: Vector Floorplan Generation via a Diffusion Model With Discrete and Continuous Denoising, *Mohammad Amin Shabani, Sepidehsadat Hosseini, Yasutaka Furukawa*
128. Localized Semantic Feature Mixers for Efficient Pedestrian Detection in Autonomous Driving, *Abdul Hannan Khan, Mohammed Shariq Nawaz, Andreas Dengel*
129. V2X-Seq: A Large-Scale Sequential Dataset for Vehicle-Infrastructure Cooperative Perception and Forecasting, *Haibao Yu, Wenxian Yang, Hongzhi Ruan, Zhenwei Yang, Yingjuan Tang, Xu Gao, Xin Hao, Yifeng Shi, Yifeng Pan, Ning Sun, Juan Song, Jirui Yuan, Ping Luo, Zaiqing Nie*
130. ViP3D: End-to-End Visual Trajectory Prediction via 3D Agent Queries, *Junru Gu, Chenxu Hu, Tianyuan Zhang, Xuanyao Chen, Yilun Wang, Yue Wang, Hang Zhao*
131. IPCC-TP: Utilizing Incremental Pearson Correlation Coefficient for Joint Multi-Agent Trajectory Prediction, *Dekai Zhu, Guangyao Zhai, Yan Di, Fabian Manhardt, Hendrik Berkemeyer, Tuan Tran, Nassir Navab, Federico Tombari, Benjamin Busam*
132. Leapfrog Diffusion Model for Stochastic Trajectory Prediction, *Weibo Mao, Chenxin Xu, Qi Zhu, Siheng Chen, Yanfeng Wang*
133. DeFeeNet: Consecutive 3D Human Motion Prediction With Deviation Feedback, *Xiaoning Sun, Huaijiang Sun, Bin Li, Dong Wei, Weiqing Li, Jianfeng Lu*
134. Self-Correctable and Adaptable Inference for Generalizable Human Pose Estimation, *Zhehan Kan, Shuoshuo Chen, Ce Zhang, Yushun Tang, Zhihai He*
135. ReDirTrans: Latent-to-Latent Translation for Gaze and Head Redirection, *Shiwei Jin, Zhen Wang, Lei Wang, Ning Bi, Truong Nguyen*
136. Feature Shrinkage Pyramid for Camouflaged Object Detection With Transformers, *Zhou Huang, Hang Dai, Tian-Zhu Xiang, Shuo Wang, Huai-Xin Chen, Jie Qin, Huan Xiong*
137. OVTrack: Open-Vocabulary Multiple Object Tracking, *Siyuan Li, Tobias Fischer, Lei Ke, Henghui Ding, Martin Danelljan, Fisher Yu*
138. GaitGCI: Generative Counterfactual Intervention for Gait Recognition, *Huanzhang Dou, Pengyi Zhang, Wei Su, Yunlong Yu, Yining Lin, Xi Li*
139. Multi-Label Compound Expression Recognition: C-EXPR Database & Network, *Dimitrios Kollias*
140. Blemish-Aware and Progressive Face Retouching With Limited Paired Data, *Lianxin Xie, Wen Xue, Zhen Xu, Si Wu, Zhiwen Yu, Hau San Wong*
141. High-Fidelity and Freely Controllable Talking Head Video Generation, *Yue Gao, Yuan Zhou, Jinglu Wang, Xiao Li, Xiang Ming, Yan Lu*
142. 3Mformer: Multi-Order Multi-Mode Transformer for Skeletal Action Recognition, *Lei Wang, Piotr Koniusz*
143. UDE: A Unified Driving Engine for Human Motion Generation, *Zixiang Zhou, Baoyuan Wang*
144. Data-Driven Feature Tracking for Event Cameras, *Nico Messikommer, Carter Fang, Mathias Gehrig, Davide Scaramuzza*
145. MoStGAN-V: Video Generation With Temporal Motion Styles, *Xiaoqian Shen, Xiang Li, Mohamed Elhoseiny*
146. Two-Stage Co-Segmentation Network Based on Discriminative Representation for Recovering Human Mesh From Videos, *Boyang Zhang, Kehua Ma, Suping Wu, Zhixiang Yuan*

147. Joint Appearance and Motion Learning for Efficient Rolling Shutter Correction, *Bin Fan, Yuxin Mao, Yuchao Dai, Zhexiong Wan, Qi Liu*
148. Extracting Motion and Appearance via Inter-Frame Attention for Efficient Video Frame Interpolation, *Guozhen Zhang, Yuhan Zhu, Haonan Wang, Youxin Chen, Gangshan Wu, Limin Wang*
149. Deep Stereo Video Inpainting, *Zhiliang Wu, Changchang Sun, Hanyu Xuan, Yan Yan*
150. Burstformer: Burst Image Restoration and Enhancement Transformer, *Akshay Dudhane, Syed Waqas Zamir, Salman Khan, Fahad Shahbaz Khan, Ming-Hsuan Yang*
151. Blur Interpolation Transformer for Real-World Motion From Blur, *Zhihang Zhong, Mingdeng Cao, Xiang Ji, Yinqiang Zheng, Imari Sato*
152. HDR Imaging With Spatially Varying Signal-to-Noise Ratios, *Yiheng Chi, Xingguang Zhang, Stanley H. Chan*
153. Light Source Separation and Intrinsic Image Decomposition Under AC Illumination, *Yusaku Yoshida, Ryo Kawahara, Takahiro Okabe*
154. Physics-Guided ISO-Dependent Sensor Noise Modeling for Extreme Low-Light Photography, *Yue Cao, Ming Liu, Shuai Liu, Xiaotao Wang, Lei Lei, Wangmeng Zuo*
155. Neumann Network With Recursive Kernels for Single Image Defocus Deblurring, *Yuhui Quan, Zicong Wu, Hui Ji*
156. UMat: Uncertainty-Aware Single Image High Resolution Material Capture, *Carlos Rodriguez-Pardo, Henar Domínguez-Elvira, David Pascual-Hernández, Elena Garces*
157. SMAE: Few-Shot Learning for HDR Deghosting With Saturation-Aware Masked Autoencoders, *Qingsen Yan, Song Zhang, Weiye Chen, Hao Tang, Yu Zhu, Jinqiu Sun, Luc Van Gool, Yanning Zhang*
158. Curricular Contrastive Regularization for Physics-Aware Single Image Dehazing, *Yu Zheng, Jiahui Zhan, Shengfeng He, Junyu Dong, Yong Du*
159. Patch-Craft Self-Supervised Training for Correlated Image Denoising, *Gregory Vaksman, Michael Elad*
160. Spectral Enhanced Rectangle Transformer for Hyperspectral Image Denoising, *Miaoyu Li, Ji Liu, Ying Fu, Yulun Zhang, Dejing Dou*
161. All-in-One Image Restoration for Unknown Degradations Using Adaptive Discriminative Filters for Specific Degradations, *Dongwon Park, Byung Hyun Lee, Se Young Chun*
162. Ingredient-Oriented Multi-Degradation Learning for Image Restoration, *Jinghao Zhang, Jie Huang, Mingde Yao, Zizheng Yang, Hu Yu, Man Zhou, Feng Zhao*
163. CR-FIQA: Face Image Quality Assessment by Learning Sample Relative Classifiability, *Fadi Boutros, Meiling Fang, Marcel Klemm, Biying Fu, Naser Damer*
164. Re-IQA: Unsupervised Learning for Image Quality Assessment in the Wild, *Avinab Saha, Sandeep Mishra, Alan C. Bovik*
165. Toward Accurate Post-Training Quantization for Image Super Resolution, *Zhijun Tu, Jie Hu, Hanting Chen, Yunhe Wang*
166. Learning Steerable Function for Efficient Image Resampling, *Jiacheng Li, Chang Chen, Wei Huang, Zhiqiang Lang, Fenglong Song, Youliang Yan, Zhiwei Xiong*
167. ABCD: Arbitrary Bitwise Coefficient for De-Quantization, *Woo Kyoung Han, Byeonghun Lee, Sang Hyun Park, Kyong Hwan Jin*
168. Efficient Frequency Domain-Based Transformers for High-Quality Image Deblurring, *Lingshun Kong, Jiangxin Dong, Jianjun Ge, Mingqiang Li, Jinshan Pan*
169. Learning a Sparse Transformer Network for Effective Image Deraining, *Xiang Chen, Hao Li, Mingqiang Li, Jinshan Pan*
170. CDDFuse: Correlation-Driven Dual-Branch Feature Decomposition for Multi-Modality Image Fusion, *Zixiang Zhao, Haowen Bai, Jianshe Zhang, Yulun Zhang, Shuang Xu, Zudi Lin, Radu Timofte, Luc Van Gool*
171. PCT-Net: Full Resolution Image Harmonization Using Pixel-Wise Color Transformations, *Julian Jorge Andrade Guerreiro, Mitsuru Nakazawa, Björn Stenger*
172. Semi-Supervised Parametric Real-World Image Harmonization, *Ke Wang, Michaël Gharbi, He Zhang, Zhihao Xia, Eli Shechtman*
173. Towards Robust Tampered Text Detection in Document Image: New Dataset and New Solution, *Chenfan Qu, Chongyu Liu, Yuliang Liu, Xinhong Chen, Dezhi Peng, Fengjun Guo, Lianwen Jin*
174. QuantArt: Quantizing Image Style Transfer Towards High Visual Fidelity, *Siyu Huang, Jie An, Donglai Wei, Jiebo Luo, Hanspeter Pfister*
175. Deep Curvilinear Editing: Commutative and Nonlinear Image Manipulation for Pretrained Deep Generative Model, *Takehiro Aoshima, Takashi Matsubara*
176. Person Image Synthesis via Denoising Diffusion Model, *Ankan Kumar Bhunia, Salman Khan, Hisham Cholakkal, Rao Muhammad Anwer, Jorma Laaksonen, Mubarak Shah, Fahad Shahbaz Khan*
177. Disentangling Writer and Character Styles for Handwriting Generation, *Gang Dai, Yifan Zhang, Qingfeng Wang, Qing Du, Zhuliang Yu, Zhuoman Liu, Shuangping Huang*
178. NoisyTwins: Class-Consistent and Diverse Image Generation Through StyleGANs, *Harsh Rangwani, Lavish Bansal, Kartik Sharma, Tejan Karmali, Varun Jampani, R. Venkatesh Babu*
179. High-Fidelity Guided Image Synthesis With Latent Diffusion Models, *Jaskirat Singh, Stephen Gould, Liang Zheng*
180. Imagic: Text-Based Real Image Editing With Diffusion Models, *Bahjat Kawar, Shiran Zada, Oran Lang, Omer Tov, Huiwen Chang, Tali Dekel, Inbar Mosseri, Michal Irani*
181. PosterLayout: A New Benchmark and Approach for Content-Aware Visual-Textual Presentation Layout, *Hsiao Yuan Hsu, Xiangteng He, Yuxin Peng, Hao Kong, Qing Zhang*
182. SINE: Single Image Editing With Text-to-Image Diffusion Models, *Zhixing Zhang, Ligong Han, Arnab Ghosh, Dimitris N. Metaxas, Jian Ren*
183. NULL-Text Inversion for Editing Real Images Using Guided Diffusion Models, *Ron Mokady, Amir Hertz, Kfir Aberman, Yael Pritch, Daniel Cohen-Or*
184. Diffusion Art or Digital Forgery? Investigating Data Replication in Diffusion Models, *Gowthami Somepalli, Vasu Singla, Micah Goldblum, Jonas Geiping, Tom Goldstein*
185. Parallel Diffusion Models of Operator and Image for Blind Inverse Problems, *Hyungjin Chung, Jeongsol Kim, Sehui Kim, Jong Chul Ye*
186. Unite and Conquer: Plug & Play Multi-Modal Synthesis Using Diffusion Models, *Nithin Gopalakrishnan Nair, Wele Gedara Chaminda Bandara, Vishal M. Patel*
187. Collaborative Diffusion for Multi-Modal Face Generation and Editing, *Ziqi Huang, Kelvin C.K. Chan, Yuming Jiang, Ziwei Liu*

188. Diffusion Video Autoencoders: Toward Temporally Consistent Face Video Editing via Disentangled Video Encoding, *Gyeongman Kim, Hajin Shim, Hyunsu Kim, Yunje Choi, Junho Kim, Eunho Yang*
189. NVTC: Nonlinear Vector Transform Coding, *Runsen Feng, Zongyu Guo, Weiping Li, Zhibo Chen*
190. Motion Information Propagation for Neural Video Compression, *Linfeng Qi, Jiahao Li, Bin Li, Houqiang Li, Yan Lu*
191. A Dynamic Multi-Scale Voxel Flow Network for Video Prediction, *Xiaotao Hu, Zhewei Huang, Ailin Huang, Jun Xu, Shuchang Zhou*
192. Towards Scalable Neural Representation for Diverse Videos, *Bo He, Xitong Yang, Hanyu Wang, Zuxuan Wu, Hao Chen, Shuaiyi Huang, Yixuan Ren, Ser-Nam Lim, Abhinav Shrivastava*
193. DINER: Disorder-Invariant Implicit Neural Representation, *Shaowen Xie, Hao Zhu, Zhen Liu, Qi Zhang, You Zhou, Xun Cao, Zhan Ma*
194. SCConv: Spatial and Channel Reconstruction Convolution for Feature Redundancy, *Jiafeng Li, Ying Wen, Lianghua He*
195. DeepMAD: Mathematical Architecture Design for Deep Convolutional Neural Network, *Xuan Shen, Yaohua Wang, Ming Lin, Yilun Huang, Hao Tang, Xiuyu Sun, Yanzhi Wang*
196. Optimization-Inspired Cross-Attention Transformer for Compressive Sensing, *Jiechong Song, Chong Mou, Shiqi Wang, Siwei Ma, Jian Zhang*
197. Neighborhood Attention Transformer, *Ali Hassani, Steven Walton, Jiachen Li, Shen Li, Humphrey Shi*
198. Making Vision Transformers Efficient From a Token Sparsification View, *Shuning Chang, Pichao Wang, Ming Lin, Fan Wang, David Junhao Zhang, Rong Jin, Mike Zheng Shou*
199. Towards Efficient Use of Multi-Scale Features in Transformer-Based Object Detectors, *Gongjie Zhang, Zhipeng Luo, Zichen Tian, Jingyi Zhang, Xiaoqin Zhang, Shijian Lu*
200. Neuralizer: General Neuroimage Analysis Without Re-Training, *Steffen Czolbe, Adrian V. Dalca*
201. Learning Partial Correlation Based Deep Visual Representation for Image Classification, *Saimunur Rahman, Piotr Koniusz, Lei Wang, Luping Zhou, Peyman Moghadam, Changming Sun*
202. Understanding Masked Image Modeling via Learning Occlusion Invariant Feature, *Xiangwen Kong, Xiangyu Zhang*
203. MixMAE: Mixed and Masked Autoencoder for Efficient Pretraining of Hierarchical Vision Transformers, *Jihao Liu, Xin Huang, Jinliang Zheng, Yu Liu, Hongsheng Li*
204. Adaptive Graph Convolutional Subspace Clustering, *Lai Wei, Zhengwei Chen, Jun Yin, Changming Zhu, Rigui Zhou, Jin Liu*
205. Deep Learning of Partial Graph Matching via Differentiable Top-K, *Runzhong Wang, Ziao Guo, Shaofei Jiang, Xiaokang Yang, Junchi Yan*
206. DynamicDet: A Unified Dynamic Architecture for Object Detection, *Zhihao Lin, Yongtao Wang, Jinhe Zhang, Xiaojie Chu*
207. IS-GGT: Iterative Scene Graph Generation With Generative Transformers, *Sanjoy Kundu, Sathyanarayanan N. Aakur*
208. Fast Contextual Scene Graph Generation With Unbiased Context Augmentation, *Tianlei Jin, Fangtai Guo, Qiwei Meng, Shiqiang Zhu, Xiangming Xi, Wen Wang, Zonghao Mu, Wei Song*
209. Masked Video Distillation: Rethinking Masked Feature Modeling for Self-Supervised Video Representation Learning, *Rui Wang, Dongdong Chen, Zuxuan Wu, Yinpeng Chen, Xiyang Dai, Mengchen Liu, Lu Yuan, Yu-Gang Jiang*
210. MED-VT: Multiscale Encoder-Decoder Video Transformer With Application to Object Segmentation, *Rezaul Karim, He Zhao, Richard P. Wildes, Mennatullah Siam*
211. MOVES: Manipulated Objects in Video Enable Segmentation, *Richard E. L. Higgins, David F. Fouhey*
212. InstMove: Instance Motion for Object-Centric Video Segmentation, *Qihao Liu, Junfeng Wu, Yi Jiang, Xiang Bai, Alan L. Yuille, Song Bai*
213. ZBS: Zero-Shot Background Subtraction via Instance-Level Background Modeling and Foreground Selection, *Yongqi An, Xu Zhao, Tao Yu, Haiyun Guo, Chaoyang Zhao, Ming Tang, Jinqiao Wang*
214. Feature Aggregated Queries for Transformer-Based Video Object Detectors, *Yiming Cui*
215. Context-Aware Relative Object Queries to Unify Video Instance and Panoptic Segmentation, *Anwesa Choudhuri, Girish Chowdhary, Alexander G. Schwing*
216. Selective Structured State-Spaces for Long-Form Video Understanding, *Jue Wang, Wentao Zhu, Pichao Wang, Xiang Yu, Linda Liu, Mohamed Omar, Raffay Hamid*
217. Relational Space-Time Query in Long-Form Videos, *Xitong Yang, Fu-Jen Chu, Matt Feiszli, Raghav Goyal, Lorenzo Torresani, Du Tran*
218. Novel-View Acoustic Synthesis, *Changan Chen, Alexander Richard, Roman Shapovalov, Vamsi Krishna Ithapu, Natalia Neverova, Kristen Grauman, Andrea Vedaldi*
219. Learning Audio-Visual Source Localization via False Negative Aware Contrastive Learning, *Weixuan Sun, Jiayi Zhang, Jianyuan Wang, Zheyuan Liu, Yiran Zhong, Tianpeng Feng, Yandong Guo, Yanhao Zhang, Nick Barnes*
220. Sound to Visual Scene Generation by Audio-to-Visual Latent Alignment, *Kim Sung-Bin, Arda Senocak, Hyunwoo Ha, Andrew Owens, Tae-Hyun Oh*
221. CASP-Net: Rethinking Video Saliency Prediction From an Audio-Visual Consistency Perceptual Perspective, *Junwen Xiong, Ganglai Wang, Peng Zhang, Wei Huang, Yufei Zha, Guangtao Zhai*
222. Decompose More and Aggregate Better: Two Closer Looks at Frequency Representation Learning for Human Motion Prediction, *Xuehao Gao, Shaoyi Du, Yang Wu, Yang Yang*
223. TempSAL – Uncovering Temporal Information for Deep Saliency Prediction, *Bahar Aydemir, Ludo Hoffstetter, Tong Zhang, Mathieu Salzmann, Sabine Süssstrunk*
224. Prompt-Guided Zero-Shot Anomaly Action Recognition Using Pretrained Deep Skeleton Features, *Fumiaki Sato, Ryo Hachiuma, Taiki Sekii*
225. MMG-Ego4D: Multimodal Generalization in Egocentric Action Recognition, *Xinyu Gong, Sreyas Mohan, Naina Dhingra, Jean-Charles Bazin, Yilei Li, Zhangyang Wang, Rakesh Ranjan*
226. Active Exploration of Multimodal Complementarity for Few-Shot Action Recognition, *Yuyang Wanyan, Xiaoshan Yang, Chaofan Chen, Changsheng Xu*
227. Reducing the Label Bias for Timestamp Supervised Temporal Action Segmentation, *Kaiyuan Liu, Yunheng Li, Shenglan Liu, Chenwei Tan, Zihang Shao*
228. Soft-Landing Strategy for Alleviating the Task Discrepancy Problem in Temporal Action Localization Tasks, *Hyolim Kang, Hanjung Kim, Joungbin An, Minsu Cho, Seon Joo Kim*

229. Iterative Proposal Refinement for Weakly-Supervised Video Grounding, *Meng Cao, Fangyun Wei, Can Xu, Xiubo Geng, Long Chen, Can Zhang, Yuexian Zou, Tao Shen, Daxin Jiang*
230. Movies2Scenes: Using Movie Metadata to Learn Scene Representation, *Shixing Chen, Chun-Hao Liu, Xiang Hao, Xiaohan Nie, Maxim Arap, Raffay Hamid*
231. Fine-Tuned CLIP Models Are Efficient Video Learners, *Hanoona Rasheed, Muhammad Uzair Khattak, Muhammad Maaz, Salman Khan, Fahad Shahbaz Khan*
232. Revisiting Temporal Modeling for CLIP-Based Image-to-Video Knowledge Transferring, *Ruyang Liu, Jingjia Huang, Ge Li, Jiashi Feng, Xinglong Wu, Thomas H. Li*
233. VoP: Text-Video Co-Operative Prompt Tuning for Cross-Modal Retrieval, *Siteng Huang, Biao Gong, Yulin Pan, Jianwen Jiang, Yiliang Lv, Yuyuan Li, Donglin Wang*
234. ProTéGé: Untrimmed Pretraining for Video Temporal Grounding by Video Temporal Grounding, *Lan Wang, Gaurav Mittal, Sandra Sajeev, Ye Yu, Matthew Hall, Vishnu Naresh Boddeti, Mei Chen*
235. Learning Video Representations From Large Language Models, ✧ *Yue Zhao, Ishan Misra, Philipp Krähenbühl, Rohit Girdhar*
236. All in One: Exploring Unified Video-Language Pre-Training, *Jinpeng Wang, Yixiao Ge, Rui Yan, Yuying Ge, Kevin Qinghong Lin, Satoshi Tsutsui, Xudong Lin, Guanyu Cai, Jianping Wu, Ying Shan, Xiaohu Qie, Mike Zheng Shou*
237. High-Fidelity Generalized Emotional Talking Face Generation With Multi-Modal Emotion Space Learning, *Chao Xu, Junwei Zhu, Jiangning Zhang, Yue Han, Wenqing Chu, Ying Tai, Chengjie Wang, Zhifeng Xie, Yong Liu*
238. Bidirectional Cross-Modal Knowledge Exploration for Video Recognition With Pre-Trained Vision-Language Models, *Wenhao Wu, Xiaohan Wang, Haipeng Luo, Jingdong Wang, Yi Yang, Wanli Ouyang*
239. Decoupled Multimodal Distilling for Emotion Recognition, ✧ *Li, Yuanzhi Wang, Zhen Cui*
240. Affection: Learning Affective Explanations for Real-World Visual Data, *Panos Achlioptas, Maks Ovsjanikov, Leonidas Guibas, Sergey Tulyakov*
241. An Actor-Centric Causality Graph for Asynchronous Temporal Inference in Group Activity, *Zhao Xie, Tian Gao, Kewei Wu, Jiao Chang*
242. VLPD: Context-Aware Pedestrian Detection via Vision-Language Semantic Self-Supervision, *Mengyin Liu, Jie Jiang, Chao Zhu, Xu-Cheng Yin*
243. 3D-Aware Object Goal Navigation via Simultaneous Exploration and Identification, *Jiazhao Zhang, Liu Dai, Fanpeng Meng, Qingnan Fan, Xuelin Chen, Kai Xu, He Wang*
244. Meta-Explore: Exploratory Hierarchical Vision-and-Language Navigation Using Scene Object Spectrum Grounding, *Minyoung Hwang, Jaeyeon Jeong, Minsoo Kim, Yoonseon Oh, Songhwa Oh*
245. NaQ: Leveraging Narrations As Queries to Supervise Episodic Memory, *Santhosh Kumar Ramakrishnan, Ziad Al-Halah, Kristen Grauman*
246. EC²: Emergent Communication for Embodied Control, *Yao Mu, Shunyu Yao, Mingyu Ding, Ping Luo, Chuang Gan*
247. Abstract Visual Reasoning: An Algebraic Approach for Solving Raven's Progressive Matrices, *Jingyi Xu, Tushar Vaidya, Yufei Wu, Saket Chandra, Zhangsheng Lai, Kai Fong Ernest Chong*
248. Logical Implications for Visual Question Answering Consistency, *Sergio Tascon-Morales, Pablo Márquez-Neila, Raphael Sznitman*
249. Divide and Conquer: Answering Questions With Object Factorization and Compositional Reasoning, *Shi Chen, Qi Zhao*
250. The Dialog Must Go On: Improving Visual Dialog via Generative Self-Training, *Gi-Cheon Kang, Sungdong Kim, Jin-Hwa Kim, Donghyun Kwak, Byoung-Tak Zhang*
251. Visual-Language Prompt Tuning With Knowledge-Guided Context Optimization, *Hantao Yao, Rui Zhang, Changsheng Xu*
252. Probabilistic Prompt Learning for Dense Prediction, *Hyeongjun Kwon, Taeyong Song, Somi Jeong, Jin Kim, Jinhyun Jang, Kwanghoon Sohn*
253. Is BERT Blind? Exploring the Effect of Vision-and-Language Pretraining on Visual Language Understanding, *Morris Alper, Michael Fiman, Hadar Averbuch-Elor*
254. Seeing What You Miss: Vision-Language Pre-Training With Semantic Completion Learning, *Yatai Ji, Rongcheng Tu, Jie Jiang, Weijie Kong, Chengfei Cai, Wenzhe Zhao, Hongfa Wang, Yujia Yang, Wei Liu*
255. Affordance Grounding From Demonstration Video to Target Image, *Joya Chen, Difei Gao, Kevin Qinghong Lin, Mike Zheng Shou*
256. Leverage Interactive Affinity for Affordance Learning, *Hongchen Luo, Wei Zhai, Jing Zhang, Yang Cao, Dacheng Tao*
257. DeAR: Debiasing Vision-Language Models With Additive Residuals, *Ashish Seth, Mayur Hemani, Chirag Agarwal*
258. Images Speak in Images: A Generalist Painter for In-Context Visual Learning, *Xinlong Wang, Wen Wang, Yue Cao, Chunhua Shen, Tiejun Huang*
259. Hyperbolic Contrastive Learning for Visual Representations Beyond Objects, *Songwei Ge, Shlok Mishra, Simon Kornblith, Chun-Liang Li, David Jacobs*
260. Picture That Sketch: Photorealistic Image Generation From Abstract Sketches, *Subhadeep Koley, Ayan Kumar Bhunia, Aneeshan Sain, Pinaki Nath Chowdhury, Tao Xiang, Yi-Zhe Song*
261. GeneCIS: A Benchmark for General Conditional Image Similarity, ✧ *Sagar Vaze, Nicolas Carion, Ishan Misra*
262. Exploiting Unlabelled Photos for Stronger Fine-Grained SBIR, *Aneeshan Sain, Ayan Kumar Bhunia, Subhadeep Koley, Pinaki Nath Chowdhury, Soumitri Chattopadhyay, Tao Xiang, Yi-Zhe Song*
263. Parts2Words: Learning Joint Embedding of Point Clouds and Texts by Bidirectional Matching Between Parts and Words, *Chuan Tang, Xi Yang, Bojian Wu, Zhizhong Han, Yi Chang*
264. Detecting and Grounding Multi-Modal Media Manipulation, *Rui Shao, Tianxing Wu, Ziwei Liu*
265. Positive-Augmented Contrastive Learning for Image and Video Captioning Evaluation, ✧ *Sara Sarto, Manuele Barraco, Marcella Cornia, Lorenzo Baraldi, Rita Cucchiara*
266. Similarity Maps for Self-Training Weakly-Supervised Phrase Grounding, *Tal Shaharabany, Lior Wolf*
267. Cross-Domain Image Captioning With Discriminative Finetuning, *Roberto Dessi, Michele Bevilacqua, Eleonora Gualdoni, Nathanaël Carraz Rakotonirina, Francesca Franzon, Marco Baroni*
268. EXIF As Language: Learning Cross-Modal Associations Between Images and Camera Metadata, ✧ *Chenhao Zheng, Ayush Shrivastava, Andrew Owens*
269. Uncurated Image-Text Datasets: Shedding Light on Demographic Bias, ✧ *Noa Garcia, Yusuke Hirota, Yankun Wu, Yuta Nakashima*

270. Filtering, Distillation, and Hard Negatives for Vision-Language Pre-Training, *Filip Radenovic, Abhimanyu Dubey, Abhishek Kadian, Todor Mihaylov, Simon Vandenhende, Yash Patel, Yi Wen, Vignesh Ramanathan, Dhruv Mahajan*
271. Turning a CLIP Model Into a Scene Text Detector, *Wenwen Yu, Yuliang Liu, Wei Hua, Deqiang Jiang, Bo Ren, Xiang Bai*
272. ScanDMM: A Deep Markov Model of Scanpath Prediction for 360° Images, *Xiangjie Sui, Yuming Fang, Hanwei Zhu, Shiqi Wang, Zhou Wang*
273. CrOC: Cross-View Online Clustering for Dense Visual Representation Learning, *Thomas Stegmüller, Tim Lebailly, Behzad Bozorgtabar, Tinne Tuytelaars, Jean-Philippe Thiran*
274. PLA: Language-Driven Open-Vocabulary 3D Scene Understanding, *Runyu Ding, Jihan Yang, Chuhui Xue, Wenqing Zhang, Song Bai, Xiaojuan Qi*
275. CLIP2Scene: Towards Label-Efficient 3D Scene Understanding by CLIP, *Runnan Chen, Youquan Liu, Lingdong Kong, Xinge Zhu, Yuexin Ma, Yikang Li, Yuenan Hou, Yu Qiao, Wenping Wang*
276. CORA: Adapting CLIP for Open-Vocabulary Detection With Region Prompting and Anchor Pre-Matching, *Xiaoshi Wu, Feng Zhu, Rui Zhao, Hongsheng Li*
277. Open-Vocabulary Attribute Detection, *María A. Bravo, Sudhanshu Mittal, Simon Ging, Thomas Brox*
278. Learning to Detect and Segment for Open Vocabulary Object Detection, *Tao Wang*
279. Open-Vocabulary Semantic Segmentation With Mask-Adapted CLIP, *Feng Liang, Bichen Wu, Xiaoliang Dai, Kunpeng Li, Yinan Zhao, Hang Zhang, Peizhao Zhang, Peter Vajda, Diana Marculescu*
280. A Simple Framework for Text-Supervised Semantic Segmentation, *Muyang Yi, Quan Cui, Hao Wu, Cheng Yang, Osamu Yoshie, Hongtao Lu*
281. GAPartNet: Cross-Category Domain-Generalizable Object Perception and Manipulation via Generalizable and Actionable Parts, *Haoran Geng, Helin Xu, Chengyang Zhao, Chao Xu, Li Yi, Siyuan Huang, He Wang*
282. GeoLayoutLM: Geometric Pre-Training for Visual Information Extraction, *Chuwei Luo, Changxu Cheng, Qi Zheng, Cong Yao*
283. Self-Supervised Image-to-Point Distillation via Semantically Tolerant Contrastive Loss, *Anas Mahmoud, Jordan S. K. Hu, Tianshu Kuai, Ali Harakeh, Liam Paull, Steven L. Waslander*
284. Generative Semantic Segmentation, *Jiaqi Chen, Jiachen Lu, Xiatian Zhu, Li Zhang*
285. MISC210K: A Large-Scale Dataset for Multi-Instance Semantic Correspondence, *Yixuan Sun, Yiwen Huang, Haijing Guo, Yuzhou Zhao, Runmin Wu, Yizhou Yu, Weifeng Ge, Wenqiang Zhang*
286. MIANet: Aggregating Unbiased Instance and General Information for Few-Shot Semantic Segmentation, *Yong Yang, Qiong Chen, Yuan Feng, Tianlin Huang*
287. PACO: Parts and Attributes of Common Objects, *Vignesh Ramanathan, Anmol Kalia, Vladan Petrovic, Yi Wen, Baixue Zheng, Baishan Guo, Rui Wang, Aaron Marquez, Rama Kovvuri, Abhishek Kadian, Amir Mousavi, Yiwen Song, Abhimanyu Dubey, Dhruv Mahajan*
288. PartDistillation: Learning Parts From Instance Segmentation, *Jang Hyun Cho, Philipp Krähenbühl, Vignesh Ramanathan*
289. ACSeg: Adaptive Conceptualization for Unsupervised Semantic Segmentation, *Kehan Li, Zhennan Wang, Zesen Cheng, Runyi Yu, Yian Zhao, Guoli Song, Chang Liu, Li Yuan, Jie Chen*
290. Reliability in Semantic Segmentation: Are We on the Right Track? *Pau de Jorge, Riccardo Volpi, Philip H.S. Torr, Grégory Rogez*
291. Rethinking the Correlation in Few-Shot Segmentation: A Buoy View, *Yuan Wang, Rui Sun, Tianzhu Zhang*
292. SIM: Semantic-Aware Instance Mask Generation for Box-Supervised Instance Segmentation, *Ruihuang Li, Chenhang He, Yabin Zhang, Shuai Li, Liyi Chen, Lei Zhang*
293. Endpoints Weight Fusion for Class Incremental Semantic Segmentation, *Jia-Wen Xiao, Chang-Bin Zhang, Jie Kang Feng, Xialei Liu, Joost van de Weijer, Ming-Ming Cheng*
294. Incremter: Transformer for Class-Incremental Semantic Segmentation With Knowledge Distillation Focusing on Old Class, *Chao Shang, Hongliang Li, Fanman Meng, Qingbo Wu, Heqian Qiu, Lanxiao Wang*
295. Continuous Pseudo-Label Rectified Domain Adaptive Semantic Segmentation With Implicit Neural Representations, *Rui Gong, Qin Wang, Martin Danelljan, Dengxin Dai, Luc Van Gool*
296. Revisiting Weak-to-Strong Consistency in Semi-Supervised Semantic Segmentation, *Lihe Yang, Lei Qi, Litong Feng, Wayne Zhang, Yinghuan Shi*
297. Discriminative Co-Saliency and Background Mining Transformer for Co-Salient Object Detection, *Long Li, Junwei Han, Ni Zhang, Nian Liu, Salman Khan, Hisham Cholakkal, Rao Muhammad Anwer, Fahad Shahbaz Khan*
298. Texture-Guided Saliency Distilling for Unsupervised Salient Object Detection, *Huajun Zhou, Bo Qiao, Lingxiao Yang, Jianhuang Lai, Xiaohua Xie*
299. An Erudite Fine-Grained Visual Classification Model, *Dongliang Chang, Yujun Tong, Ruoyi Du, Timothy Hospedales, Yi-Zhe Song, Zhanyu Ma*
300. Dynamic Graph Learning With Content-Guided Spatial-Frequency Relation Reasoning for Deepfake Detection, *Yuan Wang, Kun Yu, Chen Chen, Xiyuan Hu, Silong Peng*
301. ScaleDet: A Scalable Multi-Dataset Object Detector, *Yanbei Chen, Manchen Wang, Abhay Mittal, Zhenlin Xu, Paolo Favaro, Joseph Tighe, Davide Modolo*
302. Multi-Centroid Task Descriptor for Dynamic Class Incremental Inference, *Tenghao Cai, Zhizhong Zhang, Xin Tan, Yanyun Qu, Guannan Jiang, Chengjie Wang, Yuan Xie*
303. Matching Is Not Enough: A Two-Stage Framework for Category-Agnostic Pose Estimation, *Min Shi, Zihao Huang, Xianzheng Ma, Xiaowei Hu, Zhiguo Cao*
304. Dynamic Coarse-To-Fine Learning for Oriented Tiny Object Detection, *Chang Xu, Jian Ding, Jinwang Wang, Wen Yang, Huai Yu, Lei Yu, Gui-Song Xia*
305. Dense Distinct Query for End-to-End Object Detection, *Shilong Zhang, Xinjiang Wang, Jiaqi Wang, Jiangmiao Pang, Chengqi Lyu, Wenwei Zhang, Ping Luo, Kai Chen*
306. Meta-Tuning Loss Functions and Data Augmentation for Few-Shot Object Detection, *Berkant Demirel, Orhun Buğra Baran, Ramazan Gokberk Cinbis*
307. One-to-Few Label Assignment for End-to-End Dense Detection, *Shuai Li, Minghan Li, Ruihuang Li, Chenhang He, Lei Zhang*
308. Test Time Adaptation With Regularized Loss for Weakly Supervised Salient Object Detection, *Olga Veksler*

309. MixTeacher: Mining Promising Labels With Mixed Scale Teacher for Semi-Supervised Object Detection, *Liang Liu, Boshen Zhang, Jiangning Zhang, Wuhao Zhang, Zhenye Gan, Guanzhong Tian, Wenbing Zhu, Yabiao Wang, Chengjie Wang*
310. Exploring Incompatible Knowledge Transfer in Few-Shot Image Generation, *Yunqing Zhao, Chao Du, Milad Abdollahzadeh, Tianyu Pang, Min Lin, Shuicheng Yan, Ngai-Man Cheung*
311. Exploring Intra-Class Variation Factors With Learnable Cluster Prompts for Semi-Supervised Image Synthesis, *Yunfei Zhang, Xiaoyang Huo, Tianyi Chen, Si Wu, Hau San Wong*
312. A Soma Segmentation Benchmark in Full Adult Fly Brain, *Xiaoyu Liu, Bo Hu, Mingxing Li, Wei Huang, Yueyi Zhang, Zhiwei Xiong*
313. SDC-UDA: Volumetric Unsupervised Domain Adaptation Framework for Slice-Direction Continuous Cross-Modality Medical Image Segmentation, *Hyungseob Shin, Hyeongyu Kim, Sewon Kim, Yohan Jun, Taejoon Eo, Dosik Hwang*
314. Label-Free Liver Tumor Segmentation, *Qixin Hu, Yixiong Chen, Junfei Xiao, Shuwen Sun, Jieneng Chen, Alan L. Yuille, Zongwei Zhou*
315. Interactive and Explainable Region-Guided Radiology Report Generation, *Tim Tanida, Philip Müller, Georgios Kaissis, Daniel Rueckert*
316. A Loopback Network for Explainable Microvascular Invasion Classification, *Shengxuming Zhang, Tianqi Shi, Yang Jiang, Xiuming Zhang, Jie Lei, Zunlei Feng, Mingli Song*
317. Task-Specific Fine-Tuning via Variational Information Bottleneck for Weakly-Supervised Pathology Whole Slide Image Classification, *Honglin Li, Chenglu Zhu, Yunlong Zhang, Yuxuan Sun, Zhongyi Shui, Wenwei Kuang, Sunyi Zheng, Lin Yang*
318. YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors, *Chien-Yao Wang, Alexey Bochkovskiy, Hong-Yuan Mark Liao*
319. ✨ Two-Way Multi-Label Loss, *Takumi Kobayashi*
320. Teaching Matters: Investigating the Role of Supervision in Vision Transformers, *Matthew Walmer, Saksham Suri, Kamal Gupta, Abhinav Shrivastava*
321. Label Information Bottleneck for Label Enhancement, *Qinghai Zheng, Jihua Zhu, Haoyu Tang*
322. Glocal Energy-Based Learning for Few-Shot Open-Set Recognition, *Haoyu Wang, Guansong Pang, Peng Wang, Lei Zhang, Wei Wei, Yanning Zhang*
323. Noisy Correspondence Learning With Meta Similarity Correction, *Haochen Han, Kaiyao Miao, Qinghua Zheng, Minnan Luo*
324. Hubs and Hyperspheres: Reducing Hubness and Improving Transductive Few-Shot Learning With Hyperspherical Embeddings, *Daniel J. Trosten, Riddhi Chakraborty, Sigurd Løkse, Kristoffer Knutsen Wickstrøm, Robert Jenssen, Michael C. Kampffmeyer*
325. Coreset Sampling From Open-Set for Fine-Grained Self-Supervised Learning, *Sungnyun Kim, Sangmin Bae, Se-Young Yun*
326. Boosting Semi-Supervised Learning by Exploiting All Unlabeled Data, *Yuhao Chen, Xin Tan, Borui Zhao, Zhaowei Chen, Renjie Song, Jiajun Liang, Xuequan Lu*
327. Trade-Off Between Robustness and Accuracy of Vision Transformers, *Yanxi Li, Chang Xu*
328. Exploring and Utilizing Pattern Imbalance, *Shibin Mei, Chenglong Zhao, Shengchao Yuan, Bingbing Ni*
329. Dynamic Conceptual Contrastive Learning for Generalized Category Discovery, *Nan Pu, Zhun Zhong, Nicu Sebe*
330. Towards Better Decision Forests: Forest Alternating Optimization, *Miguel Á. Carreira-Perpiñán, Magzhan Gabidolla, Arman Zharmagambetov*
331. Learning Debaised Representations via Conditional Attribute Interpolation, *Yi-Kai Zhang, Qi-Wei Wang, De-Chuan Zhan, Han-Jia Ye*
332. On the Pitfall of Mixup for Uncertainty Calibration, *Deng-Bao Wang, Lanqing Li, Peilin Zhao, Pheng-Ann Heng, Min-Ling Zhang*
333. Class Relationship Embedded Learning for Source-Free Unsupervised Domain Adaptation, *Yixin Zhang, Zilei Wang, Weinan He*
334. FeatureBooster: Boosting Feature Descriptors With a Lightweight Neural Network, *Xinjiang Wang, Zeyu Liu, Yu Hu, Wei Xi, Wenxian Yu, Danping Zou*
335. Guiding Pseudo-Labels With Uncertainty Estimation for Source-Free Unsupervised Domain Adaptation, *Mattia Litrico, Alessio Del Bue, Pietro Morerio*
336. Divide and Adapt: Active Domain Adaptation via Customized ✨ Learning, *Duojun Huang, Jichang Li, Weikai Chen, Junshi Huang, Zhenhua Chai, Guanbin Li*
337. Understanding and Constructing Latent Modality Structures in Multi-Modal Representation Learning, *Qian Jiang, Changyou Chen, Han Zhao, Liqun Chen, Qing Ping, Son Dinh Tran, Yi Xu, Belinda Zeng, Trishul Chilimbi*
338. Deep Factorized Metric Learning, *Chengkun Wang, Wenzhao Zheng, Junlong Li, Jie Zhou, Jiwen Lu*
339. Meta-Causal Learning for Single Domain Generalization, *Jin Chen, Zhi Gao, Xinxiao Wu, Jiebo Luo*
340. Meta Omnium: A Benchmark for General-Purpose Learning-To-Learn, *Ondrej Bohdal, Yinbing Tian, Yongshuo Zong, Ruchika Chavhan, Da Li, Henry Gouk, Li Guo, Timothy Hospedales*
341. Robust Mean Teacher for Continual and Gradual Test-Time Adaptation, *Mario Döbler, Robert A. Marsden, Bin Yang*
342. NAR-Former: Neural Architecture Representation Learning Towards Holistic Attributes Prediction, *Yun Yi, Haokui Zhang, Wenze Hu, Nannan Wang, Xiaoyu Wang*
343. Visual Query Tuning: Towards Effective Usage of Intermediate Representations for Parameter and Memory Efficient Transfer Learning, *Cheng-Hao Tu, Zheda Mai, Wei-Lun Chao*
344. Architecture, Dataset and Model-Scale Agnostic Data-Free Meta-Learning, *Zixuan Hu, Li Shen, Zhenyi Wang, Tongliang Liu, Chun Yuan, Dacheng Tao*
345. GKEAL: Gaussian Kernel Embedded Analytic Learning for Few-Shot Class Incremental Task, *Huiping Zhuang, Zhenyu Weng, Run He, Zhiping Lin, Ziqian Zeng*
346. Mitigating Task Interference in Multi-Task Learning via Explicit Task Routing With Non-Learnable Primitives, *Chuntao Ding, Zhichao Lu, Shangguang Wang, Ran Cheng, Vishnu Naresh Boddeti*
347. Boundary Unlearning: Rapid Forgetting of Deep Networks via Shifting the Decision Boundary, *Min Chen, Weizhuo Gao, Gaoyang Liu, Kai Peng, Chen Wang*
348. Task Difficulty Aware Parameter Allocation & Regularization for Lifelong Learning, *Wenjin Wang, Yunqing Hu, Qianglong Chen, Yin Zhang*

349. Learning to Retain While Acquiring: Combating Distribution-Shift in Adversarial Data-Free Knowledge Distillation, *Gaurav Patel, Konda Reddy Mopuri, Qiang Qiu*
350. A Unified Knowledge Distillation Framework for Deep Directed Graphical Models, *Yizhuo Chen, Kaizhao Liang, Zhe Zeng, Shuochao Yao, Huajie Shao*
351. Coaching a Teachable Student, *Jimuyang Zhang, Zanming Huang, Eshed Ohn-Bar*
352. Adaptive Plasticity Improvement for Continual Learning, *Yan-Shuo Liang, Wu-Jun Li*
353. Improving Generalization of Meta-Learning With Inverted Regularization at Inner-Level, *Lianzhe Wang, Shiji Zhou, Shanghang Zhang, Xu Chu, Heng Chang, Wenwu Zhu*
354. Trainable Projected Gradient Method for Robust Fine-Tuning, *Junjiao Tian, Xiaoliang Dai, Chih-Yao Ma, Zecheng He, Yen-Cheng Liu, Zsolt Kira*
355. Imitation Learning As State Matching via Differentiable Physics, *Siwei Chen, Xiao Ma, Zhongwen Xu*
356. Improved Distribution Matching for Dataset Condensation, *Ganlong Zhao, Guanbin Li, Yipeng Qin, Yizhou Yu*
357. A General Regret Bound of Preconditioned Gradient Method for DNN Training, *Hongwei Yong, Ying Sun, Lei Zhang*
358. From Node Interaction to Hop Interaction: New Effective and Scalable Graph Learning Paradigm, *Jie Chen, Zilong Li, Yin Zhu, Junping Zhang, Jian Pu*
359. Constructing Deep Spiking Neural Networks From Artificial Neural Networks With Knowledge Distillation, *Qi Xu, Yaxin Li, Jiangrong Shen, Jian K. Liu, Huajin Tang, Gang Pan*
360. Rate Gradient Approximation Attack Threats Deep Spiking Neural Networks, *Tong Bu, Jianhao Ding, Zecheng Hao, Zhaofei Yu*
361. MobileOne: An Improved One Millisecond Mobile Backbone, *Pavan Kumar Anasosalu Vasu, James Gabriel, Jeff Zhu, Oncel Tuzel, Anurag Ranjan*
362. Understanding Masked Autoencoders via Hierarchical Latent Variable Models, *Lingjing Kong, Martin Q. Ma, Guangyi Chen, Eric P. Xing, Yuejie Chi, Louis-Philippe Morency, Kun Zhang*
363. Training Debiased Subnetworks With Contrastive Weight Pruning, *Geon Yeong Park, Sangmin Lee, Sang Wan Lee, Jong Chul Ye*
364. One-Shot Model for Mixed-Precision Quantization, *Ivan Koryakovskiy, Alexandra Yakovleva, Valentin Buchnev, Temur Isaev, Gleb Odnokikh*
365. Solving Oscillation Problem in Post-Training Quantization Through a Theoretical Perspective, *Yuexiao Ma, Huixia Li, Xiawu Zheng, Xuefeng Xiao, Rui Wang, Shilei Wen, Xin Pan, Fei Chao, Rongrong Ji*
366. Adaptive Data-Free Quantization, *Biao Qian, Yang Wang, Richang Hong, Meng Wang*
367. Learning to Generate Image Embeddings With User-Level Differential Privacy, *Zheng Xu, Maxwell Collins, Yuxiao Wang, Liviu Panait, Sewoong Oh, Sean Augenstein, Ting Liu, Florian Schroff, H. Brendan McMahan*
368. Cross-GAN Auditing: Unsupervised Identification of Attribute Level Similarities and Differences Between Pretrained Generative Models, *Matthew L. Olson, Shusen Liu, Rushil Anirudh, Jayaraman J. Thiagarajan, Peer-Timo Bremer, Weng-Keen Wong*
369. HandsOff: Labeled Dataset Generation With No Additional Human Annotations, *Austin Xu, Mariya I. Vasileva, Achal Dave, Arjun Seshadri*
370. Attribute-Preserving Face Dataset Anonymization via Latent Code Optimization, *Simone Barattin, Christos Tzelepis, Ioannis Patrass, Nicu Sebe*
371. Fake It Till You Make It: Learning Transferable Representations From Synthetic ImageNet Clones, *Mert Bülent Sarıyıldız, Karteek Alahari, Diane Larlus, Yannis Kalantidis*
372. Unbiased Multiple Instance Learning for Weakly Supervised Video Anomaly Detection, *Hui Lv, Zhongqi Yue, Qianru Sun, Bin Luo, Zhen Cui, Hanwang Zhang*
373. Multimodal Industrial Anomaly Detection via Hybrid Fusion, *Yue Wang, Jinlong Peng, Jiangning Zhang, Ran Yi, Yabiao Wang, Chengjie Wang*
374. FedSeg: Class-Heterogeneous Federated Learning for Semantic Segmentation, *Jiaxu Miao, Zongxin Yang, Leilei Fan, Yi Yang*
375. Decentralized Learning With Multi-Headed Distillation, *Andrey Zhmoginov, Mark Sandler, Nolan Miller, Gus Kristiansen, Max Vladymyrov*
376. Learning Federated Visual Prompt in Null Space for MRI Reconstruction, *Chun-Mei Feng, Bangjun Li, Xinxing Xu, Yong Liu, Huazhu Fu, Wangmeng Zuo*
377. Federated Learning With Data-Agnostic Distribution Fusion, *Jian-hui Duan, Wenzhong Li, Derun Zou, Ruichen Li, Sanglu Lu*
378. CaPriDe Learning: Confidential and Private Decentralized Learning Based on Encryption-Friendly Distillation Loss, *Nurbek Tastan, Karthik Nandakumar*
379. Multi-View Adversarial Discriminator: Mine the Non-Causal Factors for Object Detection in Unseen Domains, *Mingjun Xu, Lingyun Qin, Weijie Chen, Shiliang Pu, Lei Zhang*
380. Single Image Backdoor Inversion via Robust Smoothed Classifiers, *Mingjie Sun, Zico Kolter*
381. Effective Ambiguity Attack Against Passport-Based DNN Intellectual Property Protection Schemes Through Fully Connected Layer Substitution, *Yiming Chen, Jinyu Tian, Xiangyu Chen, Jiantao Zhou*
382. Color Backdoor: A Robust Poisoning Attack in Color Space, *Wenbo Jiang, Hongwei Li, Guowen Xu, Tianwei Zhang*
383. Adversarially Robust Neural Architecture Search for Graph Neural Networks, *Beini Xie, Heng Chang, Ziwei Zhang, Xin Wang, Daixin Wang, Zhiqiang Zhang, Rex Ying, Wenwu Zhu*
384. Minimizing Maximum Model Discrepancy for Transferable Black-Box Targeted Attacks, *Anqi Zhao, Tong Chu, Yahao Liu, Wen Li, Jingjing Li, Lixin Duan*
385. StyLess: Boosting the Transferability of Adversarial Examples, *Kaisheng Liang, Bin Xiao*
386. Improving the Transferability of Adversarial Samples by Path-Augmented Method, *Jianping Zhang, Jen-tse Huang, Wenxuan Wang, Yichen Li, Weibin Wu, Xiaosen Wang, Yuxin Su, Michael R. Lyu*
387. Feature Separation and Recalibration for Adversarial Robustness, *Woo Jae Kim, Yoonki Cho, Junsik Jung, Sung-Eui Yoon*
388. CFA: Class-Wise Calibrated Fair Adversarial Training, *Zeming Wei, Yifei Wang, Yiwen Guo, Yisen Wang*
389. Revisiting Residual Networks for Adversarial Robustness, *Shihua Huang, Zhichao Lu, Kalyanmoy Deb, Vishnu Naresh Boddeti*

390. Privacy-Preserving Adversarial Facial Features, *Zhibo Wang, He Wang, Shuaijin Jin, Wenwen Zhang, Jiahui Hu, Yan Wang, Peng Sun, Wei Yuan, Kaixin Liu, Kui Ren*
391. Edge-Aware Regional Message Passing Controller for Image Forgery Localization, *Dong Li, Jiaying Zhu, Menglu Wang, Jiawei Liu, Xueyang Fu, Zheng-Jun Zha*

1900–2100 Social: How to Negotiate Industry Offers in AI Proposal (West 118–120)

Organizer: Nicole Bannon

Abstract: Join our social event to get the tools, information, and data you need to negotiate your next offer more confidently. Some of the topics we'll cover in a 2 hour period (including 45 mins for Q&A) are: Understanding the fundamentals of compensation in tech (particularly around equity, bonus structures, etc.), data points for different levels/positions in AI, how to get over your fears of negotiating, how to decide which company / offer is right for you, how to negotiate without counter offers and without knowing "market value", how to respond to pushback from recruiters and other guilt tripping / lowballing / pressure tactics, how to avoid having an offer rescinded, how to negotiate deadline of an offer and walking through a timeline of the negotiation process for a new offer.

1900–2100 Social: Black in AI (West 223–224)

Organizers: Daniel Abidemi Ajisafe, Oluwabukola Grace Adegboro, Mennatullah Siam, Salomey Osei, Esther Odunayo Oduntan, Samson Kirk-Koffi, Nene Azu, Issam Laradji, Blaise Appolinary

Abstract: Africa has the second-largest population in the world with around 1.4 billion people as of 2022. With the increasing amount of visual data and the growing rate of its data footprint, the impact of extending Computer Vision research to solving local problems specific to Africa has become an ever-increasing need. This social event aims to bring together a unique community of people who self-identify as Black and/or from African origin or support the Black community at its first gathering in CVPR. Our main goal is to create a platform where Black researchers are comfortable meeting with other people without feeling out-of-place and to enforce a strong connection of like-minded individuals whose main or sub-goals is to empower the African community and Black Computer Vision researchers. This social, therefore, has several aims:

- Empowering Black and African origin researchers by affirming their sense of belonging to the Computer Vision community specifically in CVPR.
- Providing mentorship and guidance to young researchers from the Black and African origin community.
- Allowing both Black and African-origin researchers and their supporters/allies to gather and network within the Computer Vision community.
- Celebrating African grassroots in AI, especially in the field of Computer Vision.

1900–2100 Social: Diversity and Inclusion (West 220–222)

Organizers: Jiawei He, Katerina Fragkiadaki, Vikram V. Ramaswamy

Abstract: Diversity and inclusion are crucial for driving organizational impact in both academia and industry. To foster an environment that respects and cherishes a culture of inclusion and belonging, we propose organizing a social event that connects diversity and inclusion (D&I) initiatives with people attending CVPR. The goal of this event is to raise awareness and build connections among CVPR attendees who are interested in participating in or creating D&I initiatives. We believe that by bringing together organizations such as AI4ALL, Women in Computer Vision, CIFAR, Let's SOLVE it from Borealis AI, etc., this social event can serve as a bridge between different organizations and promote future collaborations.

1900–2100 Social: AMA with Senior Faculty and Industry Leaders (West 201)

Organizer: Yong Hee Lee

Abstract: This is a casual networking event for graduate students, faculty members, and industry professionals. There will be standing tables, light food, and drinks provided by the event sponsors. Students will have a chance to connect with fellow graduate students and mentors from various backgrounds, at a variety of stages throughout their careers.

1900–2100 Social: CV Entrepreneurship – Founders, Freelancers & Friends (West 211)

Organizers: Ankur Kalra, Sarah Andrews, Matthew Flagg

Abstract: CV Entrepreneurship: Founders, Freelancers & Friends comes to CVPR to help entrepreneurially minded members of the community find mentors, collaborators, and friends! Computer Vision is a vibrant and rapidly expanding field, and many folks are choosing to blaze their own paths outside of the corporate and academic worlds. Entrepreneurship presents its own unique mix of challenges and opportunities. Whether you're already a business owner looking for others who can relate to your experience, or you aspire to venture out on your own in some way, this is the event for you. Everyone has something to share, and something to learn. Walk away with a few new friends, inspiring ideas, some helpful resources, and enthusiasm toward your own entrepreneurial path. The goal of "Each One, Teach One" is to help members of the community develop curated connections over areas of mutual interest. Whether it's a specific research technique, how to magnify your impact, or how to accelerate your research experiments, someone at CVPR is passionate about sharing their point of view on that topic. Come to this event to meet them!

Everyone has something to share and something to learn. Walk away with a few new friends, some of whom are experts in areas you are curious about, and some of whom are curious about your areas of expertise.

Wednesday, June 21

0730–1600 Registration (West Ballroom Foyer)

0730–0900 Breakfast (West Ballrooms A–D)

0800–0830 Poster Setup (West Exhibit Hall)

0830–0900 Awards Ceremony (East Exhibit Halls A–B)

0900–1000 Plenary-Wed-AM (East Exhibit Halls A–B)

Keynote: Yejin Choi (*Univ. of Washington*)

Title: 2050: An AI Odyssey: The Dark Matter of Intelligence.

Chair: Judy Hoffman (*Georgia Institute of Technology*)

1000–1800 Exhibits (West Exhibit Hall)

- See Exhibits map for list of exhibitors.

1000–1230 Demos (West Exhibit Hall Demo Area)

- ConceptFusion: Real-Time Open-Set Multimodal Mapping, *Krishna Murthy Jatavallabhula, Alihusein Kuwajerwala, Qiao Gu, Mohd Omama, Tao Chen, Alaa Maalouf, Shuang Li, Ganesh Iyer, Soroush Saryazdi, Nikhil Keetha, Ayush Tewari, Joshua B. Tenenbaum, Celso Miguel de Melo, Madhava Krishna, Liam Paull, Florian Shkurti, Antonio Torralba*
- Advance On-Device ML Made Easy With MediaPipe From Google, *Lu Wang, Chuo-Ling Chang*
- Depth Covariance Monocular Dense Visual Odometry, *Eric Dexheimer, Andrew J. Davison*
- DepthFusion, *Gabriela Ben Melech, Scottie Fox, Diana Wofk, Estelle Aflalo, Shao-Yen Tseng, Jean Yu, Fabio Nonato, Alex Redden, Matthias Müller, Vasudev Lal*
- Diffusion Explainer: Interactive Visual Learning for Stable Diffusion, *Seongmin Lee, Ben Hoover, Hendrik Strobelt, Zijie J. Wang, Anthony Peng, Austin Wright, Kevin Li, Haekyu Park, Alex Yang, Polo Chau*
- Diffusion Illusions: Hiding Images in Plain Sight, *Ryan Burgert, Kanchana Ranasinge, Xiang Li, Michael Ryoo*
- DyLiN: Making Light Field Networks Dynamic, *Heng Yu, Joel Julin, Zoltan Adam Milacski, Koichiro Niinuma, Laszlo Jeni*
- EDGE Playground, *Jonathan Tseng, Rodrigo Castellon, Karen Liu*
- Energy-Efficient Adaptive 3D Sensing, *Brevin Tilmon, Zhanghao Sun, Sanjeev Koppal, Yicheng Wu, Georgios Evangelidis, Ramzi Zahreddine, Guru Krishnan, Sizhuo Ma, Jian Wang*
- EpiScope: Optical Separation of Reflected Components by Rotation of Polygonal Mirror, *Ryota Maeda, Shisaku Hiura*
- FoveaCam: A MEMS Mirror Enabled Foveating Camera, *Brevin Tilmon, Sanjeev Koppal*
- FRIDA: Robot Text-to-Drawing Demonstration, *Peter Schaldenbrand, Gaurav Parmar, Jim McCann, Jean Oh*
- Full-Pose Face Reenactment and Infancy-to-Seniorhood Age Transformation, *Gee-Sern Hsu, Li-Syun Hsiung, Jie-Ying Zhang, Wei-Jie Hong, Bor-Chen Liao*
- Hand Pressure Estimation From an RGB Camera, *Patrick Grady*

- Image Recognition-Based Calorie Tracker, *Pablo Behrens, Paul Bilokon, Simon Frei, Jason Lee, Ryan Perkins, Jan Peter, Laxmi Prajapat*
- Inline Microscopic 3D Shape Reconstruction, *Antensteiner Doris, Kapeller Christian, Mecca Roberto, Traxler Lukas*
- Interactive AI Art Gallery in Augmented Reality, *Andrew Mendez, Ilter Canberk, Miles Miller*
- Laugh Propagation, *Bart Trzynadlowski, Steph Ng*
- LenslessPiCam: A Hardware and Software Platform for Lensless Computational Imaging With a Raspberry Pi, *Eric Bezzam, Martin Vetterli, Matthieu Simeoni*

1000–1100 Morning Break (West Exhibit Hall)

1030–1230 Poster-Wed-AM (West Exhibit Hall)

✧ - Highlight paper (check it out)

☞ - Award candidate paper (see award sessions)

1. Swept-Angle Synthetic Wavelength Interferometry, *Alankar Kotwal, Anat Levin, Ioannis Gkioulekas*
2. RefSR-NeRF: Towards High Fidelity and Super Resolution View Synthesis, *Xudong Huang, Wei Li, Jie Hu, Hanting Chen, Yunhe Wang*
3. FreeNeRF: Improving Few-Shot Neural Rendering With Free Frequency Regularization, *Jiawei Yang, Marco Pavone, Yue Wang*
4. Local-to-Global Registration for Bundle-Adjusting Neural Radiance Fields, *Yue Chen, Xingyu Chen, Xuan Wang, Qi Zhang, Yu Guo, Ying Shan, Fei Wang*
5. Nerflets: Local Radiance Fields for Efficient Structure-Aware 3D Scene Representation From 2D Supervision, *Xiaoshuai Zhang, Abhijit Kundu, Thomas Funkhouser, Leonidas Guibas, Hao Su, Kyle Genova*
6. NeRF-DS: Neural Radiance Fields for Dynamic Specular Objects, *Zhiwen Yan, Chen Li, Gim Hee Lee*
7. Grid-Guided Neural Radiance Fields for Large Urban Scenes, *Linning Xu, Yuanbo Xiangli, Sida Peng, Xingang Pan, Nanxuan Zhao, Christian Theobalt, Bo Dai, Dahua Lin*
8. Learning Neural Duplex Radiance Fields for Real-Time View Synthesis, *Ziyu Wan, Christian Richardt, Aljaž Božič, Chao Li, Vijay Rengarajan, Seonghyeon Nam, Xiaoyu Xiang, Tuotuo Li, Bo Zhu, Rakesh Ranjan, Jing Liao*
9. EditableNeRF: Editing Topologically Varying Neural Radiance Fields by Key Points, *Chengwei Zheng, Wenbin Lin, Feng Xu*
10. Real-Time Neural Light Field on Mobile Devices, *Junli Cao, Huan Wang, Pavlo Chemerys, Vladislav Shakhrai, Ju Hu, Yun Fu, Denys Makoviichuk, Sergey Tulyakov, Jian Ren*
11. StyleRF: Zero-Shot 3D Style Transfer of Neural Radiance Fields, *Kunhao Liu, Fangneng Zhan, Yiwen Chen, Jiahui Zhang, Yingchen Yu, Abdulmotaleb El Saddik, Shijian Lu, Eric P. Xing*
12. Point2Pix: Photo-Realistic Point Cloud Rendering via Neural Radiance Fields, *Tao Hu, Xiaogang Xu, Shu Liu, Jiaya Jia*
13. Pointersect: Neural Rendering With Cloud-Ray Intersection, *Jen-Hao Rick Chang, Wei-Yu Chen, Anurag Ranjan, Kwang Moo Yi, Oncel Tuzel*
14. Neural Fields Meet Explicit Geometric Representations for Inverse Rendering of Urban Scenes, *Zian Wang, Tianchang Shen, Jun Gao, Shengyu Huang, Jacob Munkberg, Jon Hasselgren, Zan Gojcic, Wenzheng Chen, Sanja Fidler*

15. DANI-Net: Uncalibrated Photometric Stereo by Differentiable Shadow Handling, Anisotropic Reflectance Modeling, and Neural Inverse Rendering, *Zongrui Li, Qian Zheng, Boxin Shi, Gang Pan, Xudong Jiang*
16. MAIR: Multi-View Attention Inverse Rendering With 3D Spatially-Varying Lighting Estimation, *JunYong Choi, SeokYeong Lee, Haesol Park, Seung-Won Jung, Ig-Jae Kim, Junghyun Cho*
17. Weakly-Supervised Single-View Image Relighting, *Renjiao Yi, Chenyang Zhu, Kai Xu*
18. Controllable Light Diffusion for Portraits, *David Futschik, Kelvin Ritland, James Vecore, Sean Fanello, Sergio Orts-Escolano, Brian Curless, Daniel Šykora, Rohit Pandey*
19. RGBD2: Generative Scene Synthesis via Incremental View Inpainting Using RGBD Diffusion Models, *Jiabao Lei, Jiapeng Tang, Kui Jia*
20. Neural Lens Modeling, *Wenqi Xian, Aljaž Božič, Noah Snaveley, Christoph Lassner*
21. RealFusion: 360° Reconstruction of Any Object From a Single Image, *Luke Melas-Kyriazi, Iro Laina, Christian Rupprecht, Andrea Vedaldi*
22. Neuralangelo: High-Fidelity Neural Surface Reconstruction, *Zhaoshuo Li, Thomas Müller, Alex Evans, Russell H. Taylor, Mathias Unberath, Ming-Yu Liu, Chen-Hsuan Lin*
23. PermutoSDF: Fast Multi-View Reconstruction With Implicit Surfaces Using Permutohedral Lattices, *Radu Alexandru Rosu, Sven Behnke*
24. NeuDA: Neural Deformable Anchor for High-Fidelity Implicit Surface Reconstruction, *Bowen Cai, Jinchi Huang, Rongfei Jia, Chengfei Lv, Huan Fu*
25. NEF: Neural Edge Fields for 3D Parametric Curve Reconstruction From Multi-View Images, *Yunfan Ye, Renjiao Yi, Zhirui Gao, Chenyang Zhu, Zhiping Cai, Kai Xu*
26. NeuralField-LDM: Scene Generation With Hierarchical Latent Diffusion Models, *Seung Wook Kim, Bradley Brown, Kangxue Yin, Karsten Kreis, Katja Schwarz, Daiqing Li, Robin Rombach, Antonio Torralba, Sanja Fidler*
27. SinGRAF: Learning a 3D Generative Radiance Field for a Single Scene, *Minjung Son, Jeong Joon Park, Leonidas Guibas, Gordon Wetzstein*
28. Painting 3D Nature in 2D: View Synthesis of Natural Scenes From a Single Semantic Mask, *Shangzhan Zhang, Sida Peng, Tianrun Chen, Linzhan Mou, Haotong Lin, Kaicheng Yu, Yiyi Liao, Xiaowei Zhou*
29. Quantitative Manipulation of Custom Attributes on 3D-Aware Image Synthesis, *Hoseok Do, EunKyung Yoo, Taehyeong Kim, Chul Lee, Jin Young Choi*
30. NeRFInvertor: High Fidelity NeRF-GAN Inversion for Single-Shot Real Image Animation, *Yu Yin, Kamran Ghasedi, HsiangTao Wu, Jiaolong Yang, Xin Tong, Yun Fu*
31. PREIM3D: 3D Consistent Precise Image Attribute Editing From a Single Image, *Jianhui Li, Jianmin Li, Haoji Zhang, Shilong Liu, Zhengyi Wang, Zihao Xiao, Kaiwen Zheng, Jun Zhu*
32. Unsupervised 3D Shape Reconstruction by Part Retrieval and Assembly, *Xianghao Xu, Paul Guerrero, Matthew Fisher, Siddhartha Chaudhuri, Daniel Ritchie*
33. DiffSwap: High-Fidelity and Controllable Face Swapping via 3D-Aware Masked Diffusion, *Wenliang Zhao, Yongming Rao, Weikang Shi, Zuyan Liu, Jie Zhou, Jiwen Lu*
34. Fine-Grained Face Swapping via Regional GAN Inversion, *Zhian Liu, Maomao Li, Yong Zhang, Cairong Wang, Qi Zhang, Jue Wang, Yongwei Nie*
35. Logical Consistency and Greater Descriptive Power for Facial Hair Attribute Learning, *Haiyu Wu, Grace Bezdold, Aman Bhatta, Kevin W. Bowyer*
36. Learning a 3D Morphable Face Reflectance Model From Low-Cost Data, *Yuxuan Han, Zhibo Wang, Feng Xu*
37. StyleGAN Salon: Multi-View Latent Optimization for Pose-Invariant Hairstyle Transfer, *Sasikarn Khwanmuang, Pakkapon Phongthawee, Patsorn Sangkloy, Supasorn Suwajanakorn*
38. FaceLit: Neural 3D Relightable Faces, *Anurag Ranjan, Kwang Moo Yi, Jen-Hao Rick Chang, Oncel Tuzel*
39. FitMe: Deep Photorealistic 3D Morphable Model Avatars, *Alexandros Lattas, Stylianos Moschoglou, Stylianos Ploumpis, Baris Gecer, Jiankang Deng, Stefanos Zafeiriou*
40. NeuWigs: A Neural Dynamic Model for Volumetric Hair Capture and Animation, *Ziyan Wang, Giljoo Nam, Tuur Stuyck, Stephen Lombardi, Chen Cao, Jason Saragih, Michael Zollhöfer, Jessica Hodgins, Christoph Lassner*
41. SadTalker: Learning Realistic 3D Motion Coefficients for Stylized Audio-Driven Single Image Talking Face Animation, *Wenxuan Zhang, Xiaodong Cun, Xuan Wang, Yong Zhang, Xi Shen, Yu Guo, Ying Shan, Fei Wang*
42. High-Fidelity Clothed Avatar Reconstruction From a Single Image, *Tingting Liao, Xiaomei Zhang, Yuliang Xiu, Hongwei Yi, Xudong Liu, Guo-Jun Qi, Yong Zhang, Xuan Wang, Xiangyu Zhu, Zhen Lei*
43. Music-Driven Group Choreography, *Nhat Le, Thang Pham, Tuong Do, Erman Tjiputra, Quang D. Tran, Anh Nguyen*
44. Hand Avatar: Free-Pose Hand Animation and Rendering From Monocular Video, *Xingyu Chen, Baoyuan Wang, Heung-Yeung Shum*
45. Biomechanics-Guided Facial Action Unit Detection Through Force Modeling, *Zijun Cui, Chenyi Kuang, Tian Gao, Kartik Talamadupula, Qiang Ji*
46. Zero-Shot Pose Transfer for Unrigged Stylized 3D Characters, *Jiashun Wang, Xueting Li, Sifei Liu, Shalini De Mello, Orazio Gallo, Xiaolong Wang, Jan Kautz*
47. Invertible Neural Skinning, *Yash Kant, Aliaksandr Siarohin, Riza Alp Guler, Menglei Chai, Jian Ren, Sergey Tulyakov, Igor Gilitschenski*
48. BEDLAM: A Synthetic Dataset of Bodies Exhibiting Detailed ✨ Lifelike Animated Motion, *Michael J. Black, Priyanka Patel, Joachim Tesch, Jinlong Yang*
49. DIFu: Depth-Guided Implicit Function for Clothed Human Reconstruction, *Dae-Young Song, HeeKyung Lee, Jeongil Seo, Donghyeon Cho*
50. Complete 3D Human Reconstruction From a Single Incomplete Image, *Junying Wang, Jae Shin Yoon, Tuanfeng Y. Wang, Krishna Kumar Singh, Ulrich Neumann*
51. Learning Neural Volumetric Representations of Dynamic Humans in Minutes, *Chen Geng, Sida Peng, Zhen Xu, Hujun Bao, Xiaowei Zhou*
52. Marching-Primitives: Shape Abstraction From Signed Distance ✨ Function, *Weixiao Liu, Yuwei Wu, Sipu Ruan, Gregory S. Chirikjian*
53. Learning Analytical Posterior Probability for Human Mesh Recovery, *Qi Fang, Kang Chen, Yinghui Fan, Qing Shuai, Jiefeng Li, Weidong Zhang*

54. MagicPony: Learning Articulated 3D Animals in the Wild, *Shangzhe Wu, Ruining Li, Tomas Jakab, Christian Rupprecht, Andrea Vedaldi*
55. Visual-Tactile Sensing for In-Hand Object Reconstruction, *Wenqiang Xu, Zhenjun Yu, Han Xue, Ruolin Ye, Siqiong Yao, Cewu Lu*
56. Command-Driven Articulated Object Understanding and Manipulation, *Ruihang Chu, Zhengzhe Liu, Xiaoqing Ye, Xiao Tan, Xiaojuan Qi, Chi-Wing Fu, Jiaya Jia*
57. Target-Referenced Reactive Grasping for Dynamic Objects, *Jirong Liu, Ruo Zhang, Hao-Shu Fang, Minghao Gou, Hongjie Fang, Chenxi Wang, Sheng Xu, Hengxu Yan, Cewu Lu*
58. NeuralDome: A Neural Modeling Pipeline on Multi-View Human-Object Interactions, *Juze Zhang, Haimin Luo, Hongdi Yang, Xinru Xu, Qianyang Wu, Ye Shi, Jingyi Yu, Lan Xu, Jingya Wang*
59. A2J-Transformer: Anchor-to-Joint Transformer Network for 3D Interacting Hand Pose Estimation From a Single RGB Image, *Changlong Jiang, Yang Xiao, Cunlin Wu, Mingyang Zhang, Jinghong Zheng, Zhiguo Cao, Joey Tianyi Zhou*
60. TRACE: 5D Temporal Regression of Avatars With Dynamic Cameras in 3D Environments, *Yu Sun, Qian Bao, Wu Liu, Tao Mei, Michael J. Black*
61. BITE: Beyond Priors for Improved Three-D Dog Pose Estimation, *Nadine Rüegg, Shashank Tripathi, Konrad Schindler, Michael J. Black, Silvia Zuffi*
62. PoseFormerV2: Exploring Frequency Domain for Efficient and Robust 3D Human Pose Estimation, *Qitao Zhao, Ce Zheng, Mengyuan Liu, Pichao Wang, Chen Chen*
63. Global-to-Local Modeling for Video-Based 3D Human Pose and Shape Estimation, *Xiaolong Shen, Zongxin Yang, Xiaohan Wang, Jianxin Ma, Chang Zhou, Yi Yang*
64. TokenHPE: Learning Orientation Tokens for Efficient Head Pose Estimation via Transformers, *Cheng Zhang, Hai Liu, Yongjian Deng, Bochen Xie, Youfu Li*
65. GFIE: A Dataset and Baseline for Gaze-Following From 2D to 3D in Indoor Environments, *Zhengxi Hu, Yuxue Yang, Xiaolin Zhai, Dingye Yang, Bohan Zhou, Jingtai Liu*
66. Robot Structure Prior Guided Temporal Attention for Camera-to-Robot Pose Estimation From Image Sequence, *Yang Tian, Jiyao Zhang, Zekai Yin, Hao Dong*
67. Rigidity-Aware Detection for 6D Object Pose Estimation, *Yang Hai, Rui Song, Jiaojiao Li, Mathieu Salzmann, Yinlin Hu*
68. Crowd3D: Towards Hundreds of People Reconstruction From a Single Image, *Hao Wen, Jing Huang, Huili Cui, Haozhe Lin, Yu-Kun Lai, Lu Fang, Kun Li*
69. Object Pose Estimation With Statistical Guarantees: Conformal
✦ Keypoint Detection and Geometric Uncertainty Propagation, *Heng Yang, Marco Pavone*
70. expOSE: Accurate Initialization-Free Projective Factorization Using Exponential Regularization, *José Pedro Iglesias, Amanda Nilsson, Carl Olsson*
71. Neural Voting Field for Camera-Space 3D Hand Pose Estimation, *Lin Huang, Chung-Ching Lin, Kevin Lin, Lin Liang, Lijuan Wang, Junsong Yuan, Zicheng Liu*
72. Two-View Geometry Scoring Without Correspondences, *Axel Barroso-Laguna, Eric Brachmann, Victor Adrian Prisacariu, Gabriel J. Brostow, Daniyar Turmukhambetov*
73. Four-View Geometry With Unknown Radial Distortion, *Petr Hruby, Viktor Korotynskiy, Timothy Duff, Luke Oeding, Marc Pollefeys, Tomas Pajdla, Viktor Larsson*
74. BKinD-3D: Self-Supervised 3D Keypoint Discovery From Multi-View Videos, *Jennifer J. Sun, Lili Karashchuk, Amil Dravid, Serim Ryou, Sonia Fereidooni, John C. Tuthill, Aggelos Katsaggelos, Bingni W. Brunton, Georgia Gkioxari, Ann Kennedy, Yisong Yue, Pietro Perona*
75. BAAM: Monocular 3D Pose and Shape Reconstruction With Bi-Contextual Attention Module and Attention-Guided Modeling, *Hyo-Jun Lee, Hanul Kim, Su-Min Choi, Seong-Gyun Jeong, Yeong Jun Koh*
76. Multi-Object Manipulation via Object-Centric Neural Scattering Functions, *Stephen Tian, Yancheng Cai, Hong-Xing Yu, Sergey Zakharov, Katherine Liu, Adrien Gaidon, Yunzhu Li, Jiajun Wu*
77. Neural Part Priors: Learning to Optimize Part-Based Object
✦ Completion in RGB-D Scans, *Aleksei Bokhovkin, Angela Dai*
78. Panoptic Lifting for 3D Scene Understanding With Neural
✦ Fields, *Yawar Siddiqui, Lorenzo Porzi, Samuel Rota Bulò, Norman Müller, Matthias Nießner, Angela Dai, Peter Kotschieder*
79. Virtual Occlusions Through Implicit Depth, *Jamie Watson, Mohamed Sayed, Zawar Qureshi, Gabriel J. Brostow, Sara Vicente, Oisin Mac Aodha, Michael Firman*
80. Multiview Compressive Coding for 3D Reconstruction, *Chao-Yuan Wu, Justin Johnson, Jitendra Malik, Christoph Feichtenhofer, Georgia Gkioxari*
81. Behind the Scenes: Density Fields for Single View Reconstruction, *Felix Wimbauer, Nan Yang, Christian Rupprecht, Daniel Cremers*
82. VoxFormer: Sparse Voxel Transformer for Camera-Based 3D
✦ Semantic Scene Completion, *Yiming Li, Zhiding Yu, Christopher Choy, Chaowei Xiao, Jose M. Alvarez, Sanja Fidler, Chen Feng, Anima Anandkumar*
83. Renderable Neural Radiance Map for Visual Navigation, *Obin
✦ Kwon, Jeongho Park, Songhwa Oh*
84. Learning to Detect Mirrors From Videos via Dual Correspondences, *Jiaying Lin, Xin Tan, Rynson W.H. Lau*
85. Temporally Consistent Online Depth Estimation Using Point-Based Fusion, *Numair Khan, Eric Penner, Douglas Lanman, Lei Xiao*
86. Zero-Shot Dual-Lens Super-Resolution, *Ruikang Xu, Mingde Yao, Zhiwei Xiong*
87. Fully Self-Supervised Depth Estimation From Defocus Clue, *Haozhe Si, Bin Zhao, Dong Wang, Yunpeng Gao, Mulin Chen, Zhigang Wang, Xuelong Li*
88. MVImgNet: A Large-Scale Dataset of Multi-View Images, *Xianggang Yu, Mutian Xu, Yidan Zhang, Haolin Liu, Chongjie Ye, Yushuang Wu, Zizheng Yan, Chenming Zhu, Zhangyang Xiong, Tianyou Liang, Guanying Chen, Shuguang Cui, Xiaoguang Han*
89. Revisiting the Stack-Based Inverse Tone Mapping, *Ning Zhang, Yuyao Ye, Yang Zhao, Ronggang Wang*
90. Combining Implicit-Explicit View Correlation for Light Field Semantic Segmentation, *Ruixuan Cong, Da Yang, Rongshan Chen, Sizhe Wang, Zhenglong Cui, Hao Sheng*
91. 3D Spatial Multimodal Knowledge Accumulation for Scene Graph Prediction in Point Cloud, *Mingtao Feng, Haoran Hou, Liang Zhang, Zijie Wu, Yulan Guo, Ajmal Mian*

92. Role of Transients in Two-Bounce Non-Line-of-Sight Imaging, *Siddharth Somasundaram, Akshat Dave, Connor Henley, Ashok Veeraraghavan, Ramesh Raskar*
93. 3D Concept Learning and Reasoning From Multi-View Images, *Yining Hong, Chunru Lin, Yilun Du, Zhenfang Chen, Joshua B. Tenenbaum, Chuang Gan*
94. Viewpoint Equivariance for Multi-View 3D Object Detection, *Dian Chen, Jie Li, Vitor Guizilini, Rares Andrei Ambrus, Adrien Gaidon*
95. Tri-Perspective View for Vision-Based 3D Semantic Occupancy Prediction, *Yuanhui Huang, Wenzhao Zheng, Yunpeng Zhang, Jie Zhou, Jiwen Lu*
96. BEV@DC: Bird's-Eye View Assisted Training for Depth Completion, *Wending Zhou, Xu Yan, Yinghong Liao, Yuankai Lin, Jin Huang, Gangming Zhao, Shuguang Cui, Zhen Li*
97. Collaboration Helps Camera Overtake LiDAR in 3D Detection, *Yue Hu, Yifan Lu, Runsheng Xu, Weidi Xie, Siheng Chen, Yanfeng Wang*
98. Uni3D: A Unified Baseline for Multi-Dataset 3D Object Detection, *Bo Zhang, Jiakang Yuan, Botian Shi, Tao Chen, Yikang Li, Yu Qiao*
99. Towards Building Self-Aware Object Detectors via Reliable Uncertainty Quantification and Calibration, *Kemal Oksuz, Tom Joy, Puneet K. Dokania*
100. Depth Estimation From Camera Image and mmWave Radar Point Cloud, *Akash Deep Singh, Yunhao Ba, Ankur Sarker, Howard Zhang, Achuta Kadambi, Stefano Soatto, Mani Srivastava, Alex Wong*
101. SGLoc: Scene Geometry Encoding for Outdoor LiDAR Localization, *Wen Li, Shangshu Yu, Cheng Wang, Guosheng Hu, Siqi Shen, Chenglu Wen*
102. ConQueR: Query Contrast Voxel-DETR for 3D Object Detection, *✧ Benjin Zhu, Zhe Wang, Shaoshuai Shi, Hang Xu, Lanqing Hong, Hongsheng Li*
103. DeepMapping2: Self-Supervised Large-Scale LiDAR Map Optimization, *Chao Chen, Xinhao Liu, Yiming Li, Li Ding, Chen Feng*
104. Towards Unsupervised Object Detection From LiDAR Point Clouds, *Lunjun Zhang, Anqi Joyce Yang, Yuwen Xiong, Sergio Casas, Bin Yang, Mengye Ren, Raquel Urtasun*
105. MoDAR: Using Motion Forecasting for 3D Object Detection in Point Cloud Sequences, *Yingwei Li, Charles R. Qi, Yin Zhou, Chenxi Liu, Dragomir Anguelov*
106. Hidden Gems: 4D Radar Scene Flow Learning Using Cross-✧ Modal Supervision, *Fangqiang Ding, Andras Palffy, Dariu M. Gavrilă, Chris Xiaoxuan Lu*
107. Instant Domain Augmentation for LiDAR Semantic Segmentation, *Kwonyoung Ryu, Soonmin Hwang, Jaesik Park*
108. Less Is More: Reducing Task and Model Complexity for 3D Point Cloud Semantic Segmentation, *Li Li, Hubert P. H. Shum, Toby P. Breckon*
109. MarS3D: A Plug-and-Play Motion-Aware Model for Semantic Segmentation on Multi-Scan 3D Point Clouds, *Jiahui Liu, Chirui Chang, Jianhui Liu, Xiaoyang Wu, Lan Ma, Xiaojuan Qi*
110. 3D Semantic Segmentation in the Wild: Learning Generalized Models for Adverse-Condition Point Clouds, *Aoran Xiao, Jiaying Huang, Weihao Xuan, Ruijie Ren, Kangcheng Liu, Dayan Guan, Abdulmotaleb El Saddik, Shijian Lu, Eric P. Xing*
111. Novel Class Discovery for 3D Point Cloud Semantic Segmentation, *Luigi Riz, Cristiano Saltori, Elisa Ricci, Fabio Poiesi*
112. GD-MAE: Generative Decoder for MAE Pre-Training on LiDAR Point Clouds, *Honghui Yang, Tong He, Jiaheng Liu, Hua Chen, Boxi Wu, Binbin Lin, Xiaofei He, Wanli Ouyang*
113. Masked Scene Contrast: A Scalable Framework for Unsupervised 3D Representation Learning, *Xiaoyang Wu, Xin Wen, Xihui Liu, Hengshuang Zhao*
114. Open-Set Semantic Segmentation for Point Clouds via Adversarial Prototype Framework, *Jianan Li, Qiulei Dong*
115. ACL-SPC: Adaptive Closed-Loop System for Self-Supervised Point Cloud Completion, *Sangmin Hong, Mohsen Yavartanoo, Reyhaneh Neshatavar, Kyoung Mu Lee*
116. Fast Point Cloud Generation With Straight Flows, *Lemeng Wu, Dilin Wang, Chengyue Gong, Xingchao Liu, Yunyang Xiong, Rakesh Ranjan, Raghuraman Krishnamoorthi, Vikas Chandra, Qiang Liu*
117. PointVector: A Vector Representation in Point Cloud Analysis, *Xin Deng, WenYu Zhang, Qing Ding, XinMing Zhang*
118. ProxyFormer: Proxy Alignment Assisted Point Cloud Completion With Missing Part Sensitive Transformer, *Shanshan Li, Pan Gao, Xiaoyang Tan, Mingqiang Wei*
119. FAC: 3D Representation Learning via Foreground Aware Feature Contrast, *Kangcheng Liu, Aoran Xiao, Xiaoqin Zhang, Shijian Lu, Ling Shao*
120. Rethinking the Approximation Error in 3D Surface Fitting for Point Cloud Normal Estimation, *Hang Du, Xuejun Yan, Jingjing Wang, Di Xie, Shiliang Pu*
121. PointCert: Point Cloud Classification With Deterministic Certified Robustness Guarantees, *Jinghuai Zhang, Jinyuan Jia, Hongbin Liu, Neil Zhenqiang Gong*
122. Robust Multiview Point Cloud Registration With Reliable Pose Graph Initialization and History Reweighting, *Haiping Wang, Yuan Liu, Zhen Dong, Yulan Guo, Yu-Shen Liu, Wenping Wang, Bisheng Yang*
123. Visual Prompt Multi-Modal Tracking, *Jiawen Zhu, Simiao Lai, Xin Chen, Dong Wang, Huchuan Lu*
124. Progressive Neighbor Consistency Mining for Correspondence Pruning, *Xin Liu, Jufeng Yang*
125. Geometric Visual Similarity Learning in 3D Medical Image Self-Supervised Pre-Training, *Yuting He, Guanyu Yang, Rongjun Ge, Yang Chen, Jean-Louis Coatrieux, Boyu Wang, Shuo Li*
126. Unsupervised Visible-Infrared Person Re-Identification via Progressive Graph Matching and Alternate Learning, *Zesen Wu, Mang Ye*
127. Domain Generalized Stereo Matching via Hierarchical Visual Transformation, *Tianyu Chang, Xun Yang, Tianzhu Zhang, Meng Wang*
128. Unsupervised Cumulative Domain Adaptation for Foggy Scene Optical Flow, *Hanyu Zhou, Yi Chang, Wending Yan, Luxin Yan*
129. PVO: Panoptic Visual Odometry, *Weicai Ye, Xinyue Lan, Shuo Chen, Yuhang Ming, Xingyuan Yu, Hujun Bao, Zhaopeng Cui, Guofeng Zhang*
130. BAEFormer: Bi-Directional and Early Interaction Transformers for Bird's Eye View Semantic Segmentation, *Cong Pan, Yonghao He, Junran Peng, Qian Zhang, Wei Sui, Zhaoxiang Zhang*
131. Are We Ready for Vision-Centric Driving Streaming Perception? The ASAP Benchmark, *Xiaofeng Wang, Zheng Zhu, Yunpeng Zhang, Guan Huang, Yun Ye, Wenbo Xu, Ziwei Chen, Xingang Wang*

132. Visual Exemplar Driven Task-Prompting for Unified Perception in Autonomous Driving, *Xiwen Liang, Minzhe Niu, Jianhua Han, Hang Xu, Chunjing Xu, Xiaodan Liang*
133. MixSIM: A Hierarchical Framework for Mixed Reality Traffic Simulation, *Simon Suo, Kelvin Wong, Justin Xu, James Tu, Alexander Cui, Sergio Casas, Raquel Urtasun*
134. Uncovering the Missing Pattern: Unified Framework Towards Trajectory Imputation and Prediction, *Yi Xu, Armin Bazarjani, Hyung-gun Chi, Chiho Choi, Yun Fu*
135. MotionDiffuser: Controllable Multi-Agent Motion Prediction ✨ Using Diffusion, *Chiyu "Max" Jiang, Andre Cornman, Cheolho Park, Benjamin Sapp, Yin Zhou, Dragomir Anguelov*
136. Learning Human-to-Robot Handovers From Point Clouds, ✨ *Sammy Christen, Wei Yang, Claudia Pérez-D'Arpino, Otmar Hilliges, Dieter Fox, Yu-Wei Chao*
137. Phone2Proc: Bringing Robust Robots Into Our Chaotic World, *Matt Deitke, Rose Hendrix, Ali Farhadi, Kiana Ehsani, Aniruddha Kembhavi*
138. GazeNeRF: 3D-Aware Gaze Redirection With Neural Radiance Fields, *Alessandro Ruzzi, Xiangwei Shi, Xi Wang, Gengyan Li, Shalini De Mello, Hyung Jin Chang, Xucong Zhang, Otmar Hilliges*
139. Observation-Centric SORT: Rethinking SORT for Robust Multi-Object Tracking, *Jinkun Cao, Jiangmiao Pang, Xinshuo Weng, Rawal Khirodkar, Kris Kitani*
140. Autoregressive Visual Tracking, *Xing Wei, Yifan Bai, Yongchao Zheng, Dahu Shi, Yihong Gong*
141. OpenGait: Revisiting Gait Recognition Towards Better ✨ Practicality, *Chao Fan, Junhao Liang, Chuanfu Shen, Saihui Hou, Yongzhen Huang, Shiqi Yu*
142. Pose-Disentangled Contrastive Learning for Self-Supervised Facial Representation, *Yuanyuan Liu, Wenbin Wang, Yibing Zhan, Shaoze Feng, Kejun Liu, Zhe Chen*
143. Identity-Preserving Talking Face Generation With Landmark and Appearance Priors, *Weizhi Zhong, Chaowei Fang, Yinqi Cai, Pengxu Wei, Gangming Zhao, Liang Lin, Guanbin Li*
144. DF-Platter: Multi-Face Heterogeneous Deepfake Dataset, *Kartik Narayan, Harsh Agarwal, Kartik Thakral, Surbhi Mittal, Mayank Vatsa, Richa Singh*
145. Physics-Driven Diffusion Models for Impact Sound Synthesis From Videos, *Kun Su, Kaizhi Qian, Eli Shlizerman, Antonio Torralba, Chuang Gan*
146. Mofusion: A Framework for Denoising-Diffusion-Based Motion ✨ Synthesis, *Rishabh Dabral, Muhammad Hamza Mughal, Vladislav Golyanik, Christian Theobalt*
147. Adaptive Global Decay Process for Event Cameras, *Urbano Miguel Nunes, Ryad Benosman, Sio-Hoi Ieng*
148. Frame-Event Alignment and Fusion Network for High Frame Rate Tracking, *Jiqing Zhang, Yuanchen Wang, Wenxi Liu, Meng Li, Jinpeng Bai, Baocai Yin, Xin Yang*
149. Exploring Discontinuity for Video Frame Interpolation, *Sangjin Lee, Hyeonmin Lee, Chajin Shin, Hanbin Son, Sangyoun Lee*
150. AMT: All-Pairs Multi-Field Transforms for Efficient Frame Interpolation, *Zhen Li, Zuo-Liang Zhu, Ling-Hao Han, Qibin Hou, Chun-Le Guo, Ming-Ming Cheng*
151. Frame Interpolation Transformer and Uncertainty Guidance, *Markus Plack, Karlis Martins Briedis, Abdelaziz Djelouah, Matthias B. Hullin, Markus Gross, Christopher Schroers*
152. A Simple Baseline for Video Restoration With Grouped Spatial-Temporal Shift, *Dasong Li, Xiaoyu Shi, Yi Zhang, Ka Chun Cheung, Simon See, Xiaogang Wang, Hongwei Qin, Hongsheng Li*
153. Recurrent Homography Estimation Using Homography-Guided Image Warping and Focus Transformer, *Si-Yuan Cao, Runmin Zhang, Lun Luo, Beinan Yu, Zehua Sheng, Junwei Li, Hui-Liang Shen*
154. HyperCUT: Video Sequence From a Single Blurry Image Using Unsupervised Ordering, *Bang-Dang Pham, Phong Tran, Anh Tran, Cuong Pham, Rang Nguyen, Minh Hoai*
155. Indescribable Multi-Modal Spatial Evaluator, *Lingke Kong, X. Sharon Qi, Qijin Shen, Jiacheng Wang, Jingyi Zhang, Yanle Hu, Qichao Zhou*
156. Structured Kernel Estimation for Photon-Limited Deconvolution, *Yash Sanghvi, Zhiyuan Mao, Stanley H. Chan*
157. Polarized Color Image Denoising, *Zhuoxiao Li, Haiyang Jiang, Mingdeng Cao, Yinqiang Zheng*
158. Uncertainty-Aware Unsupervised Image Deblurring With Deep Residual Prior, *Xiaole Tang, Xile Zhao, Jun Liu, Jianli Wang, Yuchun Miao, Tiejong Zeng*
159. Low-Light Image Enhancement via Structure Modeling and Guidance, *Xiaogang Xu, Ruixing Wang, Jiangbo Lu*
160. Learning Sample Relationship for Exposure Correction, *Jie Huang, Feng Zhao, Man Zhou, Jie Xiao, Naishan Zheng, Kaiwen Zheng, Zhiwei Xiong*
161. Spatially Adaptive Self-Supervised Learning for Real-World Image Denoising, *Junyi Li, Zhilu Zhang, Xiaoyu Liu, Chaoyu Feng, Xiaotao Wang, Lei Lei, Wangmeng Zuo*
162. Quantum-Inspired Spectral-Spatial Pyramid Network for Hyperspectral Image Classification, *Jie Zhang, Yongshan Zhang, Yicong Zhou*
163. Generative Diffusion Prior for Unified Image Restoration and Enhancement, *Ben Fei, Zhaoyang Lyu, Liang Pan, Junzhe Zhang, Weidong Yang, Tianyue Luo, Bo Zhang, Bo Dai*
164. Ground-Truth Free Meta-Learning for Deep Compressive Sampling, *Xinran Qin, Yuhui Quan, Tongyao Pang, Hui Ji*
165. Recognizability Embedding Enhancement for Very Low-Resolution Face Recognition and Quality Estimation, *Jacky Chen Long Chai, Tiong-Sik Ng, Cheng-Yaw Low, Jaewoo Park, Andrew Beng Jin Teoh*
166. An Image Quality Assessment Dataset for Portraits, *Nicolas Chahine, Stefania Calarasanu, Davide Garcia-Civiero, Théo Cayla, Sira Ferradans, Jean Ponce*
167. Bitstream-Corrupted JPEG Images Are Restorable: Two-Stage Compensation and Alignment Framework for Image Restoration, *Wenyang Liu, Yi Wang, Kim-Hui Yap, Lap-Pui Chau*
168. Image Super-Resolution Using T-Tetromino Pixels, *Simon Grosche, Andy Regensky, Jürgen Seiler, André Kaup*
169. CUF: Continuous Upsampling Filters, *Cristina N. Vasconcelos, Cengiz Oztireli, Mark Matthews, Milad Hashemi, Kevin Swersky, Andrea Tagliasacchi*
170. OPE-SR: Orthogonal Position Encoding for Designing a Parameter-Free Upsampling Module in Arbitrary-Scale Image Super-Resolution, *Gaochao Song, Qian Sun, Luo Zhang, Ran Su, Jianfeng Shi, Ying He*
171. Implicit Diffusion Models for Continuous Super-Resolution, *Sicheng Gao, Xuhui Liu, Bohan Zeng, Sheng Xu, Yanjing Li, Xiaoyan Luo, Jianzhuang Liu, Xiantong Zhen, Baochang Zhang*

172. Pixels, Regions, and Objects: Multiple Enhancement for Salient Object Detection, *Yi Wang, Ruili Wang, Xin Fan, Tianzhu Wang, Xiangjian He*
173. VILA: Learning Image Aesthetics From User Comments With Vision-Language Pretraining, *Junjie Ke, Keren Ye, Jiahui Yu, Yonghui Wu, Peyman Milanfar, Feng Yang*
174. Image Cropping With Spatial-Aware Feature and Rank Consistency, *Chao Wang, Li Niu, Bo Zhang, Liqing Zhang*
175. B-Spline Texture Coefficients Estimator for Screen Content ✧✧ Image Super-Resolution, *Byeonghyun Pak, Jaewon Lee, Kyong Hwan Jin*
176. Delving StyleGAN Inversion for Image Editing: A Foundation Latent Space Viewpoint, *Hongyu Liu, Yibing Song, Qifeng Chen*
177. Learning Dynamic Style Kernels for Artistic Style Transfer, *Wenju Xu, Chengjiang Long, Yongwei Nie*
178. SVGformer: Representation Learning for Continuous Vector Graphics Using Transformers, *Defu Cao, Zhaowen Wang, Jose Echevarria, Yan Liu*
179. Learning Generative Structure Prior for Blind Text Image Super-Resolution, *Xiaoming Li, Wangmeng Zuo, Chen Change Loy*
180. Unsupervised Domain Adaption With Pixel-Level Discriminator for Image-Aware Layout Generation, *Chenchen Xu, Min Zhou, Tiezheng Ge, Yuning Jiang, Weiwei Xu*
181. Scaling Up GANs for Text-to-Image Synthesis, *Minguk Kang, ✧✧ Jun-Yan Zhu, Richard Zhang, Jaesik Park, Eli Shechtman, Sylvain Paris, Taesung Park*
182. ERNIE-ViLG 2.0: Improving Text-to-Image Diffusion Model With ✧✧ Knowledge-Enhanced Mixture-of-Denoising-Experts, *Zhida Feng, Zhenyu Zhang, Xintong Yu, Yewei Fang, Lanxin Li, Xuyi Chen, Yuxiang Lu, Jiaxiang Liu, Weichong Yin, Shikun Feng, Yu Sun, Li Chen, Hao Tian, Hua Wu, Haifeng Wang*
183. Inversion-Based Style Transfer With Diffusion Models, *Yuxin Zhang, Nisha Huang, Fan Tang, Haibin Huang, Chongyang Ma, Weiming Dong, Changsheng Xu*
184. Shifted Diffusion for Text-to-Image Generation, *Yufan Zhou, Bingchen Liu, Yizhe Zhu, Xiao Yang, Changyou Chen, Jinhui Xu*
185. LayoutDM: Discrete Diffusion Model for Controllable Layout Generation, *Naoto Inoue, Kotaro Kikuchi, Edgar Simo-Serra, Mayu Otani, Kota Yamaguchi*
186. Unpaired Image-to-Image Translation With Shortest Path Regularization, *Shaoan Xie, Yanwu Xu, Mingming Gong, Kun Zhang*
187. DiffCollage: Parallel Generation of Large Content With Diffusion Models, *Qinsheng Zhang, Jiaming Song, Xun Huang, Yongxin Chen, Ming-Yu Liu*
188. Wavelet Diffusion Models Are Fast and Scalable Image Generators, *Hao Phung, Quan Dao, Anh Tran*
189. MM-Diffusion: Learning Multi-Modal Diffusion Models for Joint Audio and Video Generation, *Ludan Ruan, Yiyang Ma, Huan Yang, Huiguo He, Bei Liu, Jianlong Fu, Nicholas Jing Yuan, Qin Jin, Baining Guo*
190. Adaptive Human Matting for Dynamic Videos, *Chung-Ching Lin, Jiang Wang, Kun Luo, Kevin Lin, Linjie Li, Lijuan Wang, Zicheng Liu*
191. LVQAC: Lattice Vector Quantization Coupled With Spatially Adaptive Companding for Efficient Learned Image Compression, *Xi Zhang, Xiaolin Wu*
192. Hierarchical B-Frame Video Coding Using Two-Layer CANF ✧✧ Without Motion Coding, *David Alexandre, Hsueh-Ming Hang, Wen-Hsiao Peng*
193. Towards High-Quality and Efficient Video Super-Resolution via ✧✧ Spatial-Temporal Data Overfitting, *Gen Li, Jie Ji, Minghai Qin, Wei Niu, Bin Ren, Fatemeh Afghah, Linke Guo, Xiaolong Ma*
194. HNeRV: A Hybrid Neural Representation for Videos, *Hao Chen, Matthew Gwilliam, Ser-Nam Lim, Abhinav Shrivastava*
195. Regularize Implicit Neural Representation by Itself, *Zheming Li, ✧✧ Hongxia Wang, Deyu Meng*
196. SMPConv: Self-Moving Point Representations for Continuous Convolution, *Sanghyeon Kim, Eunbyung Park*
197. Long Range Pooling for 3D Large-Scale Scene Understanding, *Xiang-Li Li, Meng-Hao Guo, Tai-Jiang Mu, Ralph R. Martin, Shi-Min Hu*
198. Progressive Random Convolutions for Single Domain Generalization, *Seokeon Choi, Debasmit Das, Sungha Choi, Seunghan Yang, Hyunsin Park, Sungrack Yun*
199. BiFormer: Vision Transformer With Bi-Level Routing Attention, *Lei Zhu, Xinjiang Wang, Zhanghan Ke, Wayne Zhang, Rynson W.H. Lau*
200. Beyond Attentive Tokens: Incorporating Token Importance and Diversity for Efficient Vision Transformers, *Sifan Long, Zhen Zhao, Jimin Pi, Shengsheng Wang, Jingdong Wang*
201. BioNet: A Biologically-Inspired Network for Face Recognition, *Pengyu Li*
202. Dual-Bridging With Adversarial Noise Generation for Domain Adaptive rPPG Estimation, *Jingda Du, Si-Qi Liu, Bochao Zhang, Pong C. Yuen*
203. On Data Scaling in Masked Image Modeling, *Zhenda Xie, Zheng Zhang, Yue Cao, Yutong Lin, Yixuan Wei, Qi Dai, Han Hu*
204. Hard Patches Mining for Masked Image Modeling, *Haochen Wang, Kaiyou Song, Junsong Fan, Yuxi Wang, Jin Xie, Zhaoxiang Zhang*
205. Evolved Part Masking for Self-Supervised Learning, *Zhanzhou Feng, Shiliang Zhang*
206. BASIS: Batch Aligned Spectral Embedding Space, *Or Streicher, Ido Cohen, Guy Gilboa*
207. OmniMAE: Single Model Masked Pretraining on Images and Videos, *Rohit Girdhar, Alaaeldin El-Nouby, Mannat Singh, Kalyan Vasudev Alwala, Armand Joulin, Ishan Misra*
208. ViTs for SITS: Vision Transformers for Satellite Image Time Series, *Michail Tarasiou, Erik Chavez, Stefanos Zafeiriou*
209. Probabilistic Debiasing of Scene Graphs, *Bashirul Azam Biswas, Qiang Ji*
210. Blind Video Deflickering by Neural Filtering With a Flawed Atlas, *Chenyang Lei, Xuanchi Ren, Zhaoxiang Zhang, Qifeng Chen*
211. SCOTCH and SODA: A Transformer Video Shadow Detection Framework, *Lihao Liu, Jean Prost, Lei Zhu, Nicolas Papadakis, Pietro Liò, Carola-Bibiane Schönlieb, Angelica I. Aviles-Rivero*
212. MAGViT: Masked Generative Video Transformer, *Lijun Yu, Yong ✧✧ Cheng, Kihyuk Sohn, José Lezama, Han Zhang, Huiwen Chang, Alexander G. Hauptmann, Ming-Hsuan Yang, Yuan Hao, Irfan Essa, Lu Jiang*
213. Improving Robustness of Semantic Segmentation to Motion-Blur Using Class-Centric Augmentation, *Aakanksha, A. N. Rajagopalan*

214. MobileVOS: Real-Time Video Object Segmentation Contrastive Learning Meets Knowledge Distillation, *Roy Miles, Mehmet Kerim Yucel, Bruno Manganelli, Albert Saà-Garriga*
215. Self-Supervised Video Forensics by Audio-Visual Anomaly ✨ Detection, *Chao Feng, Ziyang Chen, Andrew Owens*
216. Frame Flexible Network, *Yitian Zhang, Yue Bai, Chang Liu, Huan Wang, Sheng Li, Yun Fu*
217. System-Status-Aware Adaptive Network for Online Streaming Video Understanding, *Lin Geng Foo, Jia Gong, Zhipeng Fan, Jun Liu*
218. MDQE: Mining Discriminative Query Embeddings to Segment Occluded Instances on Challenging Videos, *Minghan Li, Shuai Li, Wangmeng Xiang, Lei Zhang*
219. Spatio-Temporal Pixel-Level Contrastive Learning-Based Source-Free Domain Adaptation for Video Semantic Segmentation, *Shao-Yuan Lo, Poojan Oza, Sumanth Chennupati, Alejandro Galindo, Vishal M. Patel*
220. Taming Diffusion Models for Audio-Driven Co-Speech Gesture Generation, *Lingting Zhu, Xian Liu, Xuanyu Liu, Rui Qian, Ziwei Liu, Lequan Yu*
221. Chat2Map: Efficient Scene Mapping From Multi-Ego Conversations, *Sagnik Majumder, Hao Jiang, Pierre Moulon, Ethan Henderson, Paul Calamia, Kristen Grauman, Vamsi Krishna Ithapu*
222. Audio-Visual Grouping Network for Sound Localization From Mixtures, *Shentong Mo, Yapeng Tian*
223. Language-Guided Audio-Visual Source Separation via Trimodal Consistency, *Reuben Tan, Arijit Ray, Andrea Burns, Bryan A. Plummer, Justin Salamon, Oriol Nieto, Bryan Russell, Kate Saenko*
224. Fine-Grained Audible Video Description, *Xuyang Shen, Dong Li, Jinxing Zhou, Zhen Qin, Bowen He, Xiaodong Han, Aixuan Li, Yuchao Dai, Lingpeng Kong, Meng Wang, Yu Qiao, Yiran Zhong*
225. Neural Koopman Pooling: Control-Inspired Temporal Dynamics Encoding for Skeleton-Based Action Recognition, *Xinghan Wang, Xin Xu, Yadong Mu*
226. Learning Discriminative Representations for Skeleton Based Action Recognition, *Huanyu Zhou, Qingjie Liu, Yunhong Wang*
227. Therbligs in Action: Video Understanding Through Motion Primitives, *Eadom Dessalene, Michael Maynard, Cornelia Fermüller, Yiannis Aloimonos*
228. Search-Map-Search: A Frame Selection Paradigm for Action Recognition, *Mingjun Zhao, Yakun Yu, Xiaoli Wang, Lei Yang, Di Niu*
229. Re2TAL: Rewiring Pretrained Video Backbones for Reversible Temporal Action Localization, *Chen Zhao, Shuming Liu, Karttikeya Mangalam, Bernard Ghanem*
230. Boosting Weakly-Supervised Temporal Action Localization With Text Information, *Guozhang Li, De Cheng, Xinpeng Ding, Nannan Wang, Xiaoyu Wang, Xinbo Gao*
231. Perception and Semantic Aware Regularization for Sequential Confidence Calibration, *Zhenghua Peng, Yu Luo, Tianshui Chen, Keke Xu, Shuangping Huang*
232. NewsNet: A Novel Dataset for Hierarchical Temporal Segmentation, *Haoqian Wu, Keyu Chen, Haozhe Liu, Mingchen Zhuge, Bing Li, Ruizhi Qiao, Xiujun Shu, Bei Gan, Liangsheng Xu, Bo Ren, Mengmeng Xu, Wentian Zhang, Raghavendra Ramachandra, Chia-Wen Lin, Bernard Ghanem*
233. Tell Me What Happened: Unifying Text-Guided Video Completion via Multimodal Masked Video Generation, *Tsu-Jui Fu, Licheng Yu, Ning Zhang, Cheng-Yang Fu, Jong-Chyi Su, William Yang Wang, Sean Bell*
234. Leveraging Temporal Context in Low Representational Power Regimes, *Camilo L. Fosco, SouYoung Jin, Emilie Josepshs, Aude Oliva*
235. Cap4Video: What Can Auxiliary Captions Do for Text-Video ✨ Retrieval? *Wenhao Wu, Haipeng Luo, Bo Fang, Jingdong Wang, Wanli Ouyang*
236. Vid2Seq: Large-Scale Pretraining of a Visual Language Model for Dense Video Captioning, *Antoine Yang, Arsha Nagrani, Paul Hongsuck Seo, Antoine Miech, Jordi Pont-Tuset, Ivan Laptev, Josef Sivic, Cordelia Schmid*
237. Procedure-Aware Pretraining for Instructional Video Understanding, *Honglu Zhou, Roberto Martín-Martín, Mubbasir Kapadia, Silvio Savarese, Juan Carlos Niebles*
238. VINDLU: A Recipe for Effective Video-and-Language Pretraining, *Feng Cheng, Xizi Wang, Jie Lei, David Crandall, Mohit Bansal, Gedas Bertasius*
239. Modular Memorability: Tiered Representations for Video Memorability Prediction, *Théo Dumont, Juan Segundo Hevia, Camilo L. Fosco*
240. Multivariate, Multi-Frequency and Multimodal: Rethinking Graph Neural Networks for Emotion Recognition in Conversation, *Feiyu Chen, Jie Shao, Shuyuan Zhu, Heng Tao Shen*
241. Distilling Cross-Temporal Contexts for Continuous Sign Language Recognition, *Leming Guo, Wanli Xue, Qing Guo, Bo Liu, Kaihua Zhang, Tiantian Yuan, Shengyong Chen*
242. You Need Multiple Exiting: Dynamic Early Exiting for Accelerating Unified Vision Language Model, *Shengkun Tang, Yaqing Wang, Zhenglun Kong, Tianchi Zhang, Yao Li, Caiwen Ding, Yanzhi Wang, Yi Liang, Dongkuan Xu*
243. Layout-Based Causal Inference for Object Navigation, *Sixian Zhang, Xinhang Song, Weijie Li, Yubing Bai, Xinyao Yu, Shuqiang Jiang*
244. Improving Vision-and-Language Navigation by Generating Future-View Image Semantics, *Jialu Li, Mohit Bansal*
245. A New Path: Scaling Vision-and-Language Navigation With Synthetic Instructions and Imitation Learning, *Aishwarya Kamath, Peter Anderson, Su Wang, Jing Yu Koh, Alexander Ku, Austin Waters, Yinfei Yang, Jason Baldridge, Zarana Parekh*
246. A-CAP: Anticipation Captioning With Commonsense Knowledge, *Duc Minh Vo, Quoc-An Luong, Akihiro Sugimoto, Hideki Nakayama*
247. Are Deep Neural Networks SMARTer Than Second Graders? *Anoop Cherian, Kuan-Chuan Peng, Suhas Lohit, Kevin A. Smith, Joshua B. Tenenbaum*
248. Fusing Pre-Trained Language Models With Multimodal Prompts Through Reinforcement Learning, *Youngjae Yu, Jiwan Chung, Heeseung Yun, Jack Hessel, Jae Sung Park, Ximing Lu, Rowan Zellers, Prithviraj Ammanabrolu, Ronan Le Bras, Gunhee Kim, Yejin Choi*
249. Language Adaptive Weight Generation for Multi-Task Visual Grounding, *Wei Su, Peihan Miao, Huanzhang Dou, Gaoang Wang, Liang Qiao, Zheyang Li, Xi Li*

250. From Images to Textual Prompts: Zero-Shot Visual Question Answering With Frozen Large Language Models, *Jiaxian Guo, Junnan Li, Dongxu Li, Anthony Meng Huat Tiong, Boyang Li, Dacheng Tao, Steven Hoi*
251. Diversity-Aware Meta Visual Prompting, *Qidong Huang, Xiaoyi Dong, Dongdong Chen, Weiming Zhang, Feifei Wang, Gang Hua, Nenghai Yu*
252. Hierarchical Prompt Learning for Multi-Task Learning, *Yajing Liu, Yuning Lu, Hao Liu, Yaozu An, Zhuoran Xu, Zhuokun Yao, Baofeng Zhang, Zhiwei Xiong, Chenguang Gui*
253. Task Residual for Tuning Vision-Language Models, *Tao Yu, Zhihe Lu, Xin Jin, Zhibo Chen, Xinchao Wang*
254. CREPE: Can Vision-Language Foundation Models Reason Compositionally? *Zixian Ma, Jerry Hong, Mustafa Omer Gul, Mona Gandhi, Irena Gao, Ranjay Krishna*
255. LOCATE: Localize and Transfer Object Parts for Weakly Supervised Affordance Grounding, *Gen Li, Varun Jampani, Deqing Sun, Laura Sevilla-Lara*
256. Overlooked Factors in Concept-Based Explanations: Dataset Choice, Concept Learnability, and Human Capability, *Vikram V. Ramaswamy, Sunnie S. Y. Kim, Ruth Fong, Olga Russakovsky*
257. Grounding Counterfactual Explanation of Image Classifiers to Textual Concept Space, *Siwon Kim, Jinoh Oh, Sungjin Lee, Seunghak Yu, Jaeyoung Do, Tara Taghavi*
258. GIVL: Improving Geographical Inclusivity of Vision-Language Models With Pre-Training Methods, *Da Yin, Feng Gao, Govind Thattai, Michael Johnston, Kai-Wei Chang*
259. Learning Bottleneck Concepts in Image Classification, *Bowen Wang, Liangzhi Li, Yuta Nakashima, Hajime Nagahara*
260. SceneTrilogy: On Human Scene-Sketch and Its Complementarity With Photo and Text, *Pinaki Nath Chowdhury, Ayan Kumar Bhunia, Aneeshan Sain, Subhadeep Koley, Tao Xiang, Yi-Zhe Song*
261. Context-Aware Alignment and Mutual Masking for 3D-Language Pre-Training, *Zhao Jin, Munawar Hayat, Yuwei Yang, Yulan Guo, Yinjie Lei*
262. MaskCLIP: Masked Self-Distillation Advances Contrastive Language-Image Pretraining, *Xiaoyi Dong, Jianmin Bao, Yinglin Zheng, Ting Zhang, Dongdong Chen, Hao Yang, Ming Zeng, Weiming Zhang, Lu Yuan, Dong Chen, Fang Wen, Nenghai Yu*
263. CLIPPO: Image-and-Language Understanding From Pixels Only, *Michael Tschannen, Basil Mustafa, Neil Houlsby*
264. ViLEM: Visual-Language Error Modeling for Image-Text Retrieval, *Yuxin Chen, Zongyang Ma, Ziqi Zhang, Zhongang Qi, Chunfeng Yuan, Ying Shan, Bing Li, Weiming Hu, Xiaohu Qie, Jianping Wu*
265. Non-Contrastive Learning Meets Language-Image Pre-Training, *Jinghao Zhou, Li Dong, Zhe Gan, Lijuan Wang, Furu Wei*
266. HAAV: Hierarchical Aggregation of Augmented Views for Image Captioning, *Chia-Wen Kuo, Zsolt Kira*
267. Learning Attribute and Class-Specific Representation Duet for Fine-Grained Fashion Analysis, *Yang Jiao, Yan Gao, Jingjing Meng, Jin Shang, Yi Sun*
268. Learning Instance-Level Representation for Large-Scale Multi-Modal Pretraining in E-Commerce, *Yang Jin, Yongzhi Li, Zehuan Yuan, Yadong Mu*
269. Cross-Image-Attention for Conditional Embeddings in Deep Metric Learning, *Dmytro Kotovenko, Pingchuan Ma, Timo Milbich, Björn Ommer*
270. Asymmetric Feature Fusion for Image Retrieval, *Hui Wu, Min Wang, Wengang Zhou, Zhenbo Lu, Houqiang Li*
271. Improving Zero-Shot Generalization and Robustness of Multi-Modal Models, *Yunhao Ge, Jie Ren, Andrew Gallagher, Yuxiao Wang, Ming-Hsuan Yang, Hartwig Adam, Laurent Itti, Balaji Lakshminarayanan, Jiaping Zhao*
272. Hint-Aug: Drawing Hints From Foundation Vision Transformers Towards Boosted Few-Shot Parameter-Efficient Tuning, *Zhongzhi Yu, Shang Wu, Yonggan Fu, Shun Yao Zhang, Yingyan (Celine) Lin*
273. Visual DNA: Representing and Comparing Images Using Distributions of Neuron Activations, *Benjamin Ramtoul, Matthew Gadd, Paul Newman, Daniele De Martini*
274. End-to-End 3D Dense Captioning With Vote2Cap-DETR, *Sijin Chen, Hongyuan Zhu, Xin Chen, Yinjie Lei, Gang Yu, Tao Chen*
275. Improving Table Structure Recognition With Visual-Alignment Sequential Coordinate Modeling, *Yongshuai Huang, Ning Lu, Dapeng Chen, Yibo Li, Zecheng Xie, Shenggao Zhu, Liangcai Gao, Wei Peng*
276. Region-Aware Pretraining for Open-Vocabulary Object Detection With Vision Transformers, *Dahun Kim, Anelia Angelova, Weicheng Kuo*
277. Mobile User Interface Element Detection via Adaptively Prompt Tuning, *Zhangxuan Gu, Zhuoer Xu, Haoxing Chen, Jun Lan, Changhua Meng, Weiqiang Wang*
278. Learning to Generate Text-Grounded Mask for Open-World Semantic Segmentation From Only Image-Text Pairs, *Junbum Cha, Jonghwan Mun, Byungseok Roh*
279. ZegCLIP: Towards Adapting CLIP for Zero-Shot Semantic Segmentation, *Ziqin Zhou, Yinjie Lei, Bowen Zhang, Lingqiao Liu, Yifan Liu*
280. Object-Aware Distillation Pyramid for Open-Vocabulary Object Detection, *Luting Wang, Yi Liu, Penghui Du, Zihan Ding, Yue Liao, Qiaosong Qi, Bialong Chen, Si Liu*
281. Learning Conditional Attributes for Compositional Zero-Shot Learning, *Qingsheng Wang, Lingqiao Liu, Chenchen Jing, Hao Chen, Guoqiang Liang, Peng Wang, Chunhua Shen*
282. CLIP-S⁴: Language-Guided Self-Supervised Semantic Segmentation, *Wenbin He, Suphanut Jamonnak, Liang Gou, Liu Ren*
283. StructVPR: Distill Structural Knowledge With Weighting Samples for Visual Place Recognition, *Yanqing Shen, Sanping Zhou, Jingwen Fu, Ruotong Wang, Shitao Chen, Nanning Zheng*
284. UniDAformer: Unified Domain Adaptive Panoptic Segmentation Transformer via Hierarchical Mask Calibration, *Jingyi Zhang, Jiaxing Huang, Xiaoqin Zhang, Shijian Lu*
285. Primitive Generation and Semantic-Related Alignment for Universal Zero-Shot Segmentation, *Shuting He, Henghui Ding, Wei Jiang*
286. Inferring and Leveraging Parts From Object Shape for Improving Semantic Image Synthesis, *Yuxiang Wei, Zhilong Ji, Xiaohe Wu, Jinfeng Bai, Lei Zhang, Wangmeng Zuo*
287. Composer: Bottom-Up Clustering and Compositing for Robust Part and Object Segmentation, *Ju He, Jieneng Chen, Ming-Xian Lin, Qihang Yu, Alan L. Yuille*
288. A Strong Baseline for Generalized Few-Shot Semantic Segmentation, *Sina Hajimiri, Malik Boudiaf, Ismail Ben Ayed, Jose Dolz*

289. DynaMask: Dynamic Mask Selection for Instance Segmentation, *Ruihuang Li, Chenhang He, Shuai Li, Yabin Zhang, Lei Zhang*
290. Focus on Details: Online Multi-Object Tracking With Diverse Fine-Grained Representation, *Hao Ren, Shoudong Han, Huilin Ding, Ziwen Zhang, Hongwei Wang, Faquan Wang*
291. Dynamic Focus-Aware Positional Queries for Semantic Segmentation, *Haoyu He, Jianfei Cai, Zizheng Pan, Jing Liu, Jing Zhang, Dacheng Tao, Bohan Zhuang*
292. Beyond mAP: Towards Better Evaluation of Instance Segmentation, *Rohit Jena, Lukas Zhorňak, Nehal Doiphode, Pratik Chaudhari, Vivek Buch, James Gee, Jianbo Shi*
293. Learning Orthogonal Prototypes for Generalized Few-Shot Semantic Segmentation, *Sun-Ao Liu, Yiheng Zhang, Zhaofan Qiu, Hongtao Xie, Yongdong Zhang, Ting Yao*
294. Weakly Supervised Semantic Segmentation via Adversarial Learning of Classifier and Reconstructor, *Hyeokjun Kweon, Sung-Hoon Yoon, Kuk-Jin Yoon*
295. SemiCVT: Semi-Supervised Convolutional Vision Transformer for Semantic Segmentation, *Huimin Huang, Shiao Xie, Lanfen Lin, Ruofeng Tong, Yen-Wei Chen, Yuexiang Li, Hong Wang, Yawen Huang, Yefeng Zheng*
296. Augmentation Matters: A Simple-Yet-Effective Approach to Semi-Supervised Semantic Segmentation, *Zhen Zhao, Lihe Yang, Sifan Long, Jimin Pi, Luping Zhou, Jingdong Wang*
297. The Devil Is in the Points: Weakly Semi-Supervised Instance Segmentation via Point-Guided Mask Representation, *Beomyoung Kim, Joonhyun Jeong, Dongyoon Han, Sung Ju Hwang*
298. Class-Incremental Exemplar Compression for Class-Incremental Learning, *Zilin Luo, Yaoyao Liu, Bernt Schiele, Qianru Sun*
299. Full or Weak Annotations? An Adaptive Strategy for Budget-Constrained Annotation Campaigns, *Javier Gamazo Tejero, Martin S. Zinkernagel, Sebastian Wolf, Raphael Sznitman, Pablo Márquez-Neila*
300. Learning Common Rationale to Improve Self-Supervised Representation for Fine-Grained Visual Recognition Problems, *Yangyang Shu, Anton van den Hengel, Lingqiao Liu*
301. Detection Hub: Unifying Object Detection Datasets via Query Adaptation on Language Embedding, *Lingchen Meng, Xiyang Dai, Yinpeng Chen, Pengchuan Zhang, Dongdong Chen, Mengchen Liu, Jianfeng Wang, Zuxuan Wu, Lu Yuan, Yu-Gang Jiang*
302. Self-Supervised AutoFlow, *Hsin-Ping Huang, Charles Herrmann, Junhwa Hur, Erika Lu, Kyle Sargent, Austin Stone, Ming-Hsuan Yang, Deqing Sun*
303. DETR With Additional Global Aggregation for Cross-Domain Weakly Supervised Object Detection, *Zongheng Tang, Yifan Sun, Si Liu, Yi Yang*
304. Detecting Everything in the Open World: Towards Universal Object Detection, *Zhenyu Wang, Yali Li, Xi Chen, Ser-Nam Lim, Antonio Torralba, Hengshuang Zhao, Shengjin Wang*
305. PROB: Probabilistic Objectness for Open World Object Detection, *Orr Zohar, Kuan-Chieh Wang, Serena Yeung*
306. Annealing-Based Label-Transfer Learning for Open World Object Detection, *Yuqing Ma, Hainan Li, Zhange Zhang, Jinyang Guo, Shanghang Zhang, Ruihao Gong, Xianglong Liu*
307. Learning Transformation-Predictive Representations for Detection and Description of Local Features, *Zihao Wang, Chunxu Wu, Yifei Yang, Zhen Li*
308. Bridging Precision and Confidence: A Train-Time Loss for Calibrating Object Detection, *Muhammad Akhtar Munir, Muhammad Haris Khan, Salman Khan, Fahad Shahbaz Khan*
309. 2PCNet: Two-Phase Consistency Training for Day-to-Night Unsupervised Domain Adaptive Object Detection, *Mikhail Kennerley, Jian-Gang Wang, Bharadwaj Veeravalli, Robby T. Tan*
310. Zero-Shot Generative Model Adaptation via Image-Specific Prompt Learning, *Jiayi Guo, Chaofei Wang, You Wu, Eric Zhang, Kai Wang, Xingqian Xu, Shiji Song, Humphrey Shi, Gao Huang*
311. AutoLabel: CLIP-Based Framework for Open-Set Video Domain Adaptation, *Giacomo Zara, Subhankar Roy, Paolo Rota, Elisa Ricci*
312. Bidirectional Copy-Paste for Semi-Supervised Medical Image Segmentation, *Yunhao Bai, Duowen Chen, Qingli Li, Wei Shen, Yan Wang*
313. Directional Connectivity-Based Segmentation of Medical Images, *Ziyun Yang, Sina Farsiu*
314. Ambiguous Medical Image Segmentation Using Diffusion Models, *Aimon Rahman, Jeya Maria Jose Valanarasu, Ilker Hacihaliloglu, Vishal M. Patel*
315. Sparse Multi-Modal Graph Transformer With Shared-Context Processing for Representation Learning of Giga-Pixel Images, *Ramin Nakhli, Puria Azadi Moghadam, Haoyang Mi, Hossein Farahani, Alexander Baras, Blake Gilks, Ali Bashashati*
316. METransformer: Radiology Report Generation by Transformer With Multiple Learnable Expert Tokens, *Zhanyu Wang, Lingqiao Liu, Lei Wang, Luping Zhou*
317. Towards Trustable Skin Cancer Diagnosis via Rewriting Model's Decision, *Siyuan Yan, Zhen Yu, Xuelin Zhang, Dwarikanath Mahapatra, Shekhar S. Chandra, Monika Janda, Peter Soyer, Zongyuan Ge*
318. Rethinking Out-of-Distribution (OOD) Detection: Masked Image Modeling Is All You Need, *Jingyao Li, Pengguang Chen, Zexin He, Shaozuo Yu, Shu Liu, Jiaya Jia*
319. MetaViewer: Towards a Unified Multi-View Representation, *Ren Wang, Haoliang Sun, Yuling Ma, Xiaoming Xi, Yilong Yin*
320. Deep Incomplete Multi-View Clustering With Cross-View Partial Sample and Prototype Alignment, *Jiaqi Jin, Siwei Wang, Zhibin Dong, Xinwang Liu, En Zhu*
321. RONO: Robust Discriminative Learning With Noisy Labels for 2D-3D Cross-Modal Retrieval, *Yanglin Feng, Hongyuan Zhu, Dezhong Peng, Xi Peng, Peng Hu*
322. Mind the Label Shift of Augmentation-Based Graph OOD Generalization, *Junchi Yu, Jian Liang, Ran He*
323. Zero-Shot Model Diagnosis, *Jinqi Luo, Zhaoning Wang, Chen Henry Wu, Dong Huang, Fernando De la Torre*
324. ProtoCon: Pseudo-Label Refinement via Online Clustering and Prototypical Consistency for Efficient Semi-Supervised Learning, *Islam Nassar, Munawar Hayat, Ehsan Abbasnejad, Hamid Reza Tofighi, Gholamreza Haffari*
325. Fine-Grained Classification With Noisy Labels, *Qi Wei, Lei Feng, Haoliang Sun, Ren Wang, Chenhui Guo, Yilong Yin*
326. Twin Contrastive Learning With Noisy Labels, *Zhizhong Huang, Junping Zhang, Hongming Shan*

327. RMLVQA: A Margin Loss Approach for Visual Question Answering With Language Biases, *Abhipsa Basu, Sravanti Addepalli, R. Venkatesh Babu*
328. Generative Bias for Robust Visual Question Answering, *Jae Won Cho, Dong-Jin Kim, Hyeonggon Ryu, In So Kweon*
329. On-the-Fly Category Discovery, *Ruoyi Du, Dongliang Chang, Kongming Liang, Timothy Hospedales, Yi-Zhe Song, Zhanyu Ma*
330. Co-Training 2^L Submodels for Visual Recognition, *Hugo Touvron, Matthieu Cord, Maxime Oquab, Piotr Bojanowski, Jakob Verbeek, Hervé Jégou*
331. Neural Dependencies Emerging From Learning Massive Categories, *Ruili Feng, Kecheng Zheng, Kai Zhu, Yujun Shen, Jian Zhao, Yukun Huang, Deli Zhao, Jingren Zhou, Michael Jordan, Zheng-Jun Zha*
332. MIC: Masked Image Consistency for Context-Enhanced Domain Adaptation, *Lukas Hoyer, Dengxin Dai, Haoran Wang, Luc Van Gool*
333. Towards Better Stability and Adaptability: Improve Online Self-
✧ Training for Model Adaptation in Semantic Segmentation, *Dong Zhao, Shuang Wang, Qi Zang, Dou Quan, Xiutiao Ye, Licheng Jiao*
334. DARE-GRAM: Unsupervised Domain Adaptation Regression by Aligning Inverse Gram Matrices, *Ismail Nejjar, Qin Wang, Olga Fink*
335. Equiangular Basis Vectors, *Yang Shen, Xuhao Sun, Xiu-Shen Wei*
336. Enhanced Multimodal Representation Learning With Cross-Modal KD, *Mengxi Chen, Linyu Xing, Yu Wang, Ya Zhang*
337. Decompose, Adjust, Compose: Effective Normalization by Playing With Frequency for Domain Generalization, *Sangrok Lee, Jongseong Bae, Ha Young Kim*
338. Back to the Source: Diffusion-Driven Adaptation to Test-Time Corruption, *Jin Gao, Jialing Zhang, Xihui Liu, Trevor Darrell, Evan Shelhamer, Dequan Wang*
339. Deep Frequency Filtering for Domain Generalization, *Shiqi Lin, Zhizheng Zhang, Zhipeng Huang, Yan Lu, Cuiling Lan, Peng Chu, Quanzeng You, Jiang Wang, Zicheng Liu, Amey Parulkar, Viraj Navkal, Zhibo Chen*
340. Generalizable Implicit Neural Representations via Instance
✧ Pattern Composers, *Chiheon Kim, Doyup Lee, Saehoon Kim, Minsu Cho, Wook-Shin Han*
341. Train-Once-for-All Personalization, *Hong-You Chen, Yandong Li, Yin Cui, Mingda Zhang, Wei-Lun Chao, Li Zhang*
342. Mod-Squad: Designing Mixtures of Experts As Modular Multi-Task Learners, *Zitian Chen, Yikang Shen, Mingyu Ding, Zhenfang Chen, Hengshuang Zhao, Erik G. Learned-Miller, Chuang Gan*
343. Few-Shot Class-Incremental Learning via Class-Aware Bilateral Distillation, *Linglan Zhao, Jing Lu, Yunlu Xu, Zhazhan Cheng, Dashan Guo, Yi Niu, Xiangzhong Fang*
344. Multi-Mode Online Knowledge Distillation for Self-Supervised Visual Representation Learning, *Kaiyou Song, Jin Xie, Shan Zhang, Zimeng Luo*
345. Dense Network Expansion for Class Incremental Learning, *Zhiyuan Hu, Yunsheng Li, Jiancheng Lyu, Dashan Gao, Nuno Vasconcelos*
346. Class Attention Transfer Based Knowledge Distillation, *Ziyao Guo, Haonan Yan, Hui Li, Xiaodong Lin*
347. Dealing With Cross-Task Class Discrimination in Online Continual Learning, *Yiduo Guo, Bing Liu, Dongyan Zhao*
348. Real-Time Evaluation in Online Continual Learning: A New
✧ Hope, *Yasir Ghunaim, Adel Bibi, Kumail Alhamoud, Motasem Alfarrar, Hasan Abed Al Kader Hammoud, Ameya Prabhu, Philip H.S. Torr, Bernard Ghanem*
349. DisWOT: Student Architecture Search for Distillation WithOut Training, *Peijie Dong, Lujun Li, Zimian Wei*
350. CODA-Prompt: COntinual Decomposed Attention-Based Prompting for Rehearsal-Free Continual Learning, *James Seale Smith, Leonid Karlinsky, Vyshnavi Gutta, Paola Cascante-Bonilla, Donghyun Kim, Assaf Arbelle, Rameswar Panda, Rogerio Feris, Zsolt Kira*
351. EcoTTA: Memory-Efficient Continual Test-Time Adaptation via Self-Distilled Regularization, *Junha Song, Jungsoo Lee, In So Kweon, Sungha Choi*
352. Achieving a Better Stability-Plasticity Trade-Off via Auxiliary Networks in Continual Learning, *Sanghwan Kim, Lorenzo Noci, Antonio Orvieto, Thomas Hofmann*
353. PA&DA: Jointly Sampling Path and Data for Consistent NAS, *Shun Lu, Yu Hu, Longxing Yang, Zihao Sun, Jilin Mei, Jianchao Tan, Chengru Song*
354. Accelerating Dataset Distillation via Model Augmentation, *Lei
✧ Zhang, Jie Zhang, Bowen Lei, Subhabrata Mukherjee, Xiang Pan, Bo Zhao, Caiwen Ding, Yao Li, Dongkuan Xu*
355. Multi-Agent Automated Machine Learning, *Zhaozhi Wang, Kefan Su, Jian Zhang, Huizhu Jia, Qixiang Ye, Xiaodong Xie, Zongqing Lu*
356. Transformer-Based Learned Optimization, *Erik Gärtner, Luke Metz, Mykhaylo Andriluka, C. Daniel Freeman, Cristian Sminchisescu*
357. Solving Relaxations of MAP-MRF Problems: Combinatorial In-Face Frank-Wolfe Directions, *Vladimir Kolmogorov*
358. HOTNAS: Hierarchical Optimal Transport for Neural Architecture Search, *Jiechao Yang, Yong Liu, Hongteng Xu*
359. Disentangled Representation Learning for Unsupervised Neural Quantization, *Haechan Noh, Sangeek Hyun, Woojin Jeong, Hanshin Lim, Jae-Pil Heo*
360. FFCV: Accelerating Training by Removing Data Bottlenecks, *Guillaume Leclerc, Andrew Ilyas, Logan Engstrom, Sung Min Park, Hadi Salman, Aleksander Mądry*
361. Run, Don't Walk: Chasing Higher FLOPS for Faster Neural Networks, *Jierun Chen, Shiu-hong Kao, Hao He, Weipeng Zhuo, Song Wen, Chul-Ho Lee, S.-H. Gary Chan*
362. FIANCEE: Faster Inference of Adversarial Networks via Conditional Early Exits, *Polina Karpikova, Ekaterina Radionova, Anastasia Yaschenko, Andrei Spiridonov, Leonid Kostyushko, Riccardo Fabbriatore, Aleksei Ivakhnenko*
363. Gradient-Based Uncertainty Attribution for Explainable Bayesian Deep Learning, *Hanjing Wang, Dhiraj Joshi, Shiqiang Wang, Qiang Ji*
364. How to Prevent the Continuous Damage of Noises to Model Training? *Xiaotian Yu, Yang Jiang, Tianqi Shi, Zunlei Feng, Yuexuan Wang, Mingli Song, Li Sun*
365. Genie: Show Me the Data for Quantization, *Yongkweon Jeon, Chungman Lee, Ho-young Kim*
366. OpenMix: Exploring Outlier Samples for Misclassification
✧ Detection, *Fei Zhu, Zhen Cheng, Xu-Yao Zhang, Cheng-Lin Liu*
367. Data-Free Sketch-Based Image Retrieval, *Abhra Chaudhuri, Ayan Kumar Bhunia, Yi-Zhe Song, Anjan Dutta*

- 1200-1400 Lunch** (West Ballrooms A-D)

Notes:

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.

Supported by:



- Shahira Abousamra (Stony Brook Univ.)
- Ranya Almohsen (West Virginia Univ.)
- Mohamed El Banani (Univ. of Michigan)
- Fabio Cermelli (Politecnico di Torino)
- MingFang Chang (Carnegie Mellon Univ.)
- Junwen Chen (Michigan State Univ.)
- Tianlong Chen (University of Texas at Austin)
- Xu Chen (ETH Zürich)
- Zezhou Cheng (Univ. of Massachusetts at Amherst)
- Mingyu Ding (Univ. of California at Berkeley)
- Jiahua Dong (Chinese Academy of Sciences)
- Enrico Fini (Univ. of Trento)
- Yuying Ge (Univ. of Hong Kong)
- Mariana-Iuliana Georgescu (Univ. of Bucharest)
- Shuting He (Zhejiang Univ.)
- Vincent Tao Hu (Univ. of Amsterdam)
- Jiaxing Huang (Nanyang Technological Univ.)
- José Pedro Iglesias (Chalmers Univ. of Technology)
- Yasamin Jafarian (Univ. of Minnesota)
- Yongcheng Jing (The Univ. of Sydney)
- Lei Ke (Hong Kong Univ. of Science and Technology)
- Itai Lang (Tel Aviv Univ.)
- Jungbeom Lee (Seoul National Univ.)
- Yi Li (Univ. of California at San Diego)
- Zhiheng Li (Univ. of Rochester)
- Zhuowan Li (Johns Hopkins Univ.)
- Wei Lin (Graz Univ. of Technology)
- Xudong Lin (Columbia Univ.)
- Yaoyao Liu (Johns Hopkins Univ.)
- Zhijian Liu (Massachusetts Institute of Technology)
- Shweta Mahajan (Univ. of British Columbia)
- Utkarsh Mall (Cornell Univ.)
- Chengzhi Mao (Columbia Univ.)
- Patrice Monkam (Tsinghua Univ.)
- Sida Peng (Zhejiang Univ.)
- Santhosh Kumar Ramakrishnan (Univ. of Texas at Austin)
- Vikram V. Ramaswamy (Princeton Univ.)
- Simon Reiß (Karlsruhe Institute of Technology)
- Akash Sengupta (Univ. of Cambridge)
- Bokui Shen (Stanford Univ.)
- Yujiao Shi (Australian National Univ.)
- Ivan Skorokhodov (King Abdullah Univ. of Science and Technology)
- James S. Smith (Georgia Institute of Technology)
- Guolei Sun (ETH Zürich)
- Jennifer J. Sun (California Institute of Technology)
- Xiaoxiao Sun (Australian National Univ.)
- Shixiang Tang (The Univ. of Sydney)

- Tekla Tóth (*Eötvös Loránd Univ.*)
- Ziyu Wan (*City Univ. of Hong Kong*)
- Fangyin Wei (*Princeton Univ.*)
- Sanghyun Woo (*Korea Advanced Institute of Science and Technology*)
- Taihong Xiao (*Univ. of California at Merced*)
- Gengshan Yang (*Carnegie Mellon Univ.*)
- Yiwu Zhong (*Univ. of Wisconsin at Madison*)
- Zhihang Zhong (*The Univ. of Tokyo*)

Notes:

A full-page view of a blank sheet of graph paper. The page is covered by a uniform grid of thin black lines forming small squares. There are no margins, text, or other markings on the paper.

1230–1400 Poster Switch/Setup (West Exhibit Hall)

1400–1600 Demos (West Exhibit Hall Demo Area)

- Same as morning demos (see page 25).

1400–1500 Plenary-Wed-PM (East Exhibit Halls A–B)

Panel: Vision, Language, and Creativity

Moderator: Judy Hoffman (*Georgia Institute of Technology*)

Participants:

- Aaron Hertzmann (*Adobe*)
- Michal Irani (*Weizmann Institute of Science*)
- Devi Parikh (*Georgia Institute of Technology*)
- Jason Salavon (*University of Chicago*)

1500–1600 PAMI TC Meeting (East Exhibit Halls A–B)

1600–1700 Afternoon Break (West Exhibit Hall)

1630–1830 Poster-Wed-PM (West Exhibit Hall)

✧ - Highlight paper (check it out)

☞ - Award candidate paper (see award sessions)

1. Polarimetric iToF: Measuring High-Fidelity Depth Through Scattering Media, *Daniel S. Jeon, Andréas Meuleman, Seung-Hwan Baek, Min H. Kim*
2. NeRFLix: High-Quality Neural View Synthesis by Learning a Degradation-Driven Inter-Viewpoint MiXer, *Kun Zhou, Wenbo Li, Yi Wang, Tao Hu, Nianjuan Jiang, Xiaoguang Han, Jiangbo Lu*
3. SUDS: Scalable Urban Dynamic Scenes, *Haithem Turki, Jason Y. Zhang, Francesco Ferroni, Deva Ramanan*
4. DP-NeRF: Deblurred Neural Radiance Field With Physical Scene Priors, *Dogyoon Lee, Minhyeok Lee, Chajin Shin, Sangyoun Lee*
5. DyLiN: Making Light Field Networks Dynamic, *Heng Yu, Joel Julin, Zoltán Á. Milacski, Koichiro Niinuma, László A. Jeni*
6. Multi-Space Neural Radiance Fields, *Ze-Xin Yin, Jiaxiong Qiu, Ming-Ming Cheng, Bo Ren*
7. NeRFLight: Fast and Light Neural Radiance Fields Using a Shared Feature Grid, *Fernando Rivas-Manzanique, Jorge Sierra-Acosta, Adrian Penate-Sanchez, Francesc Moreno-Noguer, Angela Ribeiro*
8. Cross-Guided Optimization of Radiance Fields With Multi-View Image Super-Resolution for High-Resolution Novel View Synthesis, *Youngho Yoon, Kuk-Jin Yoon*
9. NeuralEditor: Editing Neural Radiance Fields via Manipulating Point Clouds, *Jun-Kun Chen, Jipeng Lyu, Yu-Xiong Wang*
10. DINER: Depth-Aware Image-Based NEural Radiance Fields, *Malte Prinzler, Otmar Hilliges, Justus Thies*
11. Modernizing Old Photos Using Multiple References via Photorealistic Style Transfer, *Agus Gunawan, Soo Ye Kim, Hyeonjun Sim, Jae-Ho Lee, Munchurl Kim*
12. Efficient Map Sparsification Based on 2D and 3D Discretized Grids, *Xiaoyu Zhang, Yun-Hui Liu*
13. K-Planes: Explicit Radiance Fields in Space, Time, and Appearance, *Sara Fridovich-Keil, Giacomo Meanti, Frederik Rahbæk Warburg, Benjamin Recht, Angjoo Kanazawa*

14. I²-SDF: Intrinsic Indoor Scene Reconstruction and Editing via Raytracing in Neural SDFs, *Jingsen Zhu, Yuchi Huo, Qi Ye, Fujun Luan, Jifan Li, Dianbing Xi, Lisha Wang, Rui Tang, Wei Hua, Hujun Bao, Rui Wang*
15. Multi-View Inverse Rendering for Large-Scale Real-World Indoor Scenes, *Zhen Li, Lingli Wang, Mofang Cheng, Cihui Pan, Jiaqi Yang*
16. Inverse Rendering of Translucent Objects Using Physical and Neural Renderers, *Chenhao Li, Trung Thanh Ngo, Hajime Nagahara*
17. Accidental Light Probes, *Hong-Xing Yu, Samir Agarwala, Charles Herrmann, Richard Szeliski, Noah Snively, Jiajun Wu, Deqing Sun*
18. Humans As Light Bulbs: 3D Human Reconstruction From Thermal Reflection, *Ruoshi Liu, Carl Vondrick*
19. HumanGen: Generating Human Radiance Fields With Explicit Priors, *Suyi Jiang, Haoran Jiang, Ziyu Wang, Haimin Luo, Wenzheng Chen, Lan Xu*
20. Seeing Through the Glass: Neural 3D Reconstruction of Object Inside a Transparent Container, *Jinguang Tong, Sundaram Muthu, Fahira Afzal Maken, Chuong Nguyen, Hongdong Li*
21. 3D Shape Reconstruction of Semi-Transparent Worms, *Thomas P. Ilett, Omer Yuval, Thomas Ranner, Netta Cohen, David C. Hogg*
22. Dionysus: Recovering Scene Structures by Dividing Into Semantic Pieces, *Likang Wang, Lei Chen*
23. SparseFusion: Distilling View-Conditioned Diffusion for 3D Reconstruction, *Zhizhuo Zhou, Shubham Tulsiani*
24. PET-NeuS: Positional Encoding Tri-Planes for Neural Surfaces, *Yiqun Wang, Ivan Skorokhodov, Peter Wonka*
25. RenderDiffusion: Image Diffusion for 3D Reconstruction, Inpainting and Generation, *Titas Anciukevičius, Zexiang Xu, Matthew Fisher, Paul Henderson, Hakan Bilen, Niloy J. Mitra, Paul Guerrero*
26. Score Jacobian Chaining: Lifting Pretrained 2D Diffusion Models for 3D Generation, *Haochen Wang, Xiaodan Du, Jiahao Li, Raymond A. Yeh, Greg Shakhnarovich*
27. Infinite Photorealistic Worlds Using Procedural Generation, *Alexander Raistrick, Lahav Lipson, Zeyu Ma, Lingjie Mei, Mingzhe Wang, Yiming Zuo, Karhan Kayan, Hongyu Wen, Beining Han, Yihan Wang, Alejandro Newell, Hei Law, Ankit Goyal, Kaiyu Yang, Jia Deng*
28. Diffusion-SDF: Text-To-Shape via Voxelized Diffusion, *Muheng Li, Yueqi Duan, Jie Zhou, Jiwen Lu*
29. 3D-Aware Multi-Class Image-to-Image Translation With NeRFs, *Senmao Li, Joost van de Weijer, Yaxing Wang, Fahad Shahbaz Khan, Meiqin Liu, Jian Yang*
30. Latent-NeRF for Shape-Guided Generation of 3D Shapes and Textures, *Gal Metzer, Elad Richardson, Or Patashnik, Raja Giryes, Daniel Cohen-Or*
31. Local 3D Editing via 3D Distillation of CLIP Knowledge, *Junha Hyung, Sungwon Hwang, Daemin Kim, Hyunji Lee, Jaegul Choo*
32. ShapeTalk: A Language Dataset and Framework for 3D Shape Edits and Deformations, *Panos Achlioptas, Ian Huang, Minhyuk Sung, Sergey Tulyakov, Leonidas Guibas*
33. CoralStyleCLIP: Co-Optimized Region and Layer Selection for Image Editing, *Ambareesh Revanur, Debraj Basu, Shradha Agrawal, Dhwanit Agarwal, Deepak Pai*
34. 3D-Aware Face Swapping, *Yixuan Li, Chao Ma, Yichao Yan, Wenhan Zhu, Xiaokang Yang*

35. DCFace: Synthetic Face Generation With Dual Condition Diffusion Model, *Minchul Kim, Feng Liu, Anil Jain, Xiaoming Liu*
36. HairStep: Transfer Synthetic to Real Using Strand and Depth ✧ Maps for Single-View 3D Hair Modeling, *Yujian Zheng, Zirong Jin, Moran Li, Haibin Huang, Chongyang Ma, Shuguang Cui, Xiaoguang Han*
37. DiffuserRig: Learning Personalized Priors for Facial Appearance Editing, *Zheng Ding, Xuaner Zhang, Zhihao Xia, Lars Jebe, Zhuowen Tu, Xiuming Zhang*
38. 3D-Aware Facial Landmark Detection via Multi-View Consistent Training on Synthetic Data, *Libing Zeng, Lele Chen, Wentao Bao, Zhong Li, Yi Xu, Junsong Yuan, Nima Khademi Kalantari*
39. Parametric Implicit Face Representation for Audio-Driven Facial Reenactment, *Ricong Huang, Peiwen Lai, Yipeng Qin, Guanbin Li*
40. MEGANE: Morphable Eyeglass and Avatar Network, *Junxuan Li, Shunsuke Saito, Tomas Simon, Stephen Lombardi, Hongdong Li, Jason Saragih*
41. CodeTalker: Speech-Driven 3D Facial Animation With Discrete Motion Prior, *Jinbo Xing, Menghan Xia, Yuechen Zhang, Xiaodong Cun, Jue Wang, Tien-Tsin Wong*
42. Reconstructing Signing Avatars From Video Using Linguistic Priors, *Maria-Paola Forte, Peter Kulits, Chun-Hao P. Huang, Vasileios Choutas, Dimitrios Tzionas, Katherine J. Kuchenbecker, Michael J. Black*
43. HARP: Personalized Hand Reconstruction From a Monocular RGB Video, *Korrawe Karunratanakul, Sergey Prokudin, Otmar Hilliges, Siyu Tang*
44. OmniAvatar: Geometry-Guided Controllable 3D Head Synthesis, *Hongyi Xu, Guoxian Song, Zihang Jiang, Jianfeng Zhang, Yichun Shi, Jing Liu, Wanchun Ma, Jiashi Feng, Linjie Luo*
45. RaBit: Parametric Modeling of 3D Biped Cartoon Characters With a Topological-Consistent Dataset, *Zhongjin Luo, Shengcai Cai, Jinguo Dong, Ruibo Ming, Liangdong Qiu, Xiaohang Zhan, Xiaoguang Han*
46. Transfer4D: A Framework for Frugal Motion Capture and Deformation Transfer, *Shubh Maheshwari, Rahul Narain, Ramya Hebbalaguppe*
47. CLOTH4D: A Dataset for Clothed Human Reconstruction, *Xingxing Zou, Xintong Han, Waikeng Wong*
48. Vid2Avatar: 3D Avatar Reconstruction From Videos in the Wild via Self-Supervised Scene Decomposition, *Chen Guo, Tianjian Jiang, Xu Chen, Jie Song, Otmar Hilliges*
49. High-Fidelity 3D Human Digitization From Single 2K Resolution ✧ Images, *Sang-Hun Han, Min-Gyu Park, Ju Hong Yoon, Ju-Mi Kang, Young-Jae Park, Hae-Gon Jeon*
50. Sampling Is Matter: Point-Guided 3D Human Mesh Reconstruction, *Jeonghwan Kim, Mi-Gyeong Gwon, Hyunwoo Park, Hyukmin Kwon, Gi-Mun Um, Wonjun Kim*
51. gSDF: Geometry-Driven Signed Distance Functions for 3D Hand-Object Reconstruction, *Zerui Chen, Shizhe Chen, Cordelia Schmid, Ivan Laptev*
52. Human Body Shape Completion With Implicit Shape and Flow Learning, *Boyao Zhou, Di Meng, Jean-Sébastien Franco, Edmond Boyer*
53. ShapeClipper: Scalable 3D Shape Learning From Single-View Images via Geometric and CLIP-Based Consistency, *Zixuan Huang, Varun Jampani, Anh Thai, Yuanzhen Li, Stefan Stojanov, James M. Rehg*
54. PC²: Projection-Conditioned Point Cloud Diffusion for Single-✧ Image 3D Reconstruction, *Luke Melas-Kyriazi, Christian Ruppel, Andrea Vedaldi*
55. NIKI: Neural Inverse Kinematics With Invertible Neural Networks for 3D Human Pose and Shape Estimation, *Jiefeng Li, Siyuan Bian, Qi Liu, Jiasheng Tang, Fan Wang, Cewu Lu*
56. ARCTIC: A Dataset for Dexterous Bimanual Hand-Object Manipulation, *Zicong Fan, Omid Taheri, Dimitrios Tzionas, Muhammed Kocabas, Manuel Kaufmann, Michael J. Black, Otmar Hilliges*
57. ACR: Attention Collaboration-Based Regressor for Arbitrary Two-Hand Reconstruction, *Zhengdi Yu, Shaoli Huang, Chen Fang, Toby P. Breckon, Jue Wang*
58. MIME: Human-Aware 3D Scene Generation, *Hongwei Yi, Chun-Hao P. Huang, Shashank Tripathi, Lea Hering, Justus Thies, Michael J. Black*
59. CIMI4D: A Large Multimodal Climbing Motion Dataset Under Human-Scene Interactions, *Ming Yan, Xin Wang, Yudi Dai, Siqi Shen, Chenglu Wen, Lan Xu, Yuexin Ma, Cheng Wang*
60. Harmonious Feature Learning for Interactive Hand-Object Pose Estimation, *Zhifeng Lin, Changxing Ding, Huan Yao, Zengsheng Kuang, Shaoli Huang*
61. AssemblyHands: Towards Egocentric Activity Understanding via 3D Hand Pose Estimation, *Takehiko Ohkawa, Kun He, Fadime Sener, Tomas Hodan, Luan Tran, Cem Keskin*
62. A Characteristic Function-Based Method for Bottom-Up Human Pose Estimation, *Haoxuan Qu, Yujun Cai, Lin Geng Foo, Ajay Kumar, Jun Liu*
63. Unified Pose Sequence Modeling, *Lin Geng Foo, Tianjiao Li, Hossein Rahmani, Qihong Ke, Jun Liu*
64. Scene-Aware Egocentric 3D Human Pose Estimation, *Jian Wang, Diogo Luvizon, Weipeng Xu, Lingjie Liu, Kripasindhu Sarkar, Christian Theobalt*
65. DiffPose: Toward More Reliable 3D Pose Estimation, *Jia Gong, Lin Geng Foo, Zhipeng Fan, Qihong Ke, Hossein Rahmani, Jun Liu*
66. MammalNet: A Large-Scale Video Benchmark for Mammal Recognition and Behavior Understanding, *Jun Chen, Ming Hu, Darren J. Coker, Michael L. Berumen, Blair Costelloe, Sara Beery, Anna Rohrbach, Mohamed Elhoseiny*
67. Learning 3D-Aware Image Synthesis With Unknown Pose Distribution, *Zifan Shi, Yujun Shen, Yinghao Xu, Sida Peng, Yiyi Liao, Sheng Guo, Qifeng Chen, Dit-Yan Yeung*
68. Pose Synchronization Under Multiple Pair-Wise Relative Poses, *Yifan Sun, Qixing Huang*
69. ObjectMatch: Robust Registration Using Canonical Object Correspondences, *Can Gümel, Angela Dai, Matthias Nießner*
70. Learning Articulated Shape With Keypoint Pseudo-Labels From Web Images, *Anastasis Stathopoulos, Georgios Pavlakos, Ligong Han, Dimitris N. Metaxas*
71. Learning Correspondence Uncertainty via Differentiable Nonlinear Least Squares, *Dominik Muhle, Lukas Koestler, Krishna Murthy Jatavallabhula, Daniel Cremers*
72. ✧ Efficient Second-Order Plane Adjustment, *Lipu Zhou*
73. Learning a Depth Covariance Function, *Eric Dexheimer, Andrew J. Davison*
74. Privacy-Preserving Representations Are Not Enough: Recovering Scene Content From Camera Poses, *Kunal Chelani, Torsten Sattler, Fredrik Kahl, Zuzana Kukelova*

75. Objaverse: A Universe of Annotated 3D Objects, *Matt Deitke, Dustin Schwenk, Jordi Salvador, Luca Weihs, Oscar Michel, Eli VanderBilt, Ludwig Schmidt, Kiana Ehsani, Aniruddha Kembhavi, Ali Farhadi*
76. Omni3D: A Large Benchmark and Model for 3D Object Detection in the Wild, *Garrick Brazil, Abhinav Kumar, Julian Straub, Nikhila Ravi, Justin Johnson, Georgia Gkioxari*
77. HelixSurf: A Robust and Efficient Neural Implicit Surface Learning of Indoor Scenes With Iterative Intertwined Regularization, *Zhihao Liang, Zhangjin Huang, Changxing Ding, Kui Jia*
78. Visual Localization Using Imperfect 3D Models From the Internet, *Vojtech Panek, Zuzana Kukelova, Torsten Sattler*
79. PRISE: Demystifying Deep Lucas-Kanade With Strongly Star-
✧ Convex Constraints for Multimodel Image Alignment, *Yiqing Zhang, Xinming Huang, Ziming Zhang*
80. Scalable, Detailed and Mask-Free Universal Photometric
✧ Stereo, *Satoshi Ikehata*
81. Enhanced Stable View Synthesis, *Nishant Jain, Suryansh Kumar, Luc Van Gool*
82. End-to-End Vectorized HD-Map Construction With Piecewise Bezier Curve, *Limeng Qiao, Wenjie Ding, Xi Qiu, Chi Zhang*
83. DynamicStereo: Consistent Dynamic Depth From Stereo Videos, *Nikita Karaev, Ignacio Rocco, Benjamin Graham, Natalia Neverova, Andrea Vedaldi, Christian Rupprecht*
84. Shakes on a Plane: Unsupervised Depth Estimation From Unstabilized Photography, *Ilya Chugunov, Yuxuan Zhang, Felix Heide*
85. Gated Stereo: Joint Depth Estimation From Gated and Wide-
✧ Baseline Active Stereo Cues, *Stefanie Walz, Mario Bijelic, Andrea Ramazzina, Amanpreet Walia, Fahim Mannan, Felix Heide*
86. K3DN: Disparity-Aware Kernel Estimation for Dual-Pixel Defocus Deblurring, *Yan Yang, Liyuan Pan, Liu Liu, Miaomiao Liu*
87. HRDFuse: Monocular 360° Depth Estimation by Collaboratively Learning Holistic-With-Regional Depth Distributions, *Hao Ai, Zidong Cao, Yan-Pei Cao, Ying Shan, Lin Wang*
88. OSRT: Omnidirectional Image Super-Resolution With Distortion-Aware Transformer, *Fanghua Yu, Xintao Wang, Mingdeng Cao, Gen Li, Ying Shan, Chao Dong*
89. Co-SLAM: Joint Coordinate and Sparse Parametric Encodings for Neural Real-Time SLAM, *Hengyi Wang, Jingwen Wang, Lourdes Agapito*
90. Few-Shot Non-Line-of-Sight Imaging With Signal-Surface Collaborative Regularization, *Xintong Liu, Jianyu Wang, Leping Xiao, Xing Fu, Lingyun Qiu, Zuoqiang Shi*
91. NLOST: Non-Line-of-Sight Imaging With Transformer, *Yue Li, Jiayong Peng, Juntian Ye, Yueyi Zhang, Feihu Xu, Zhiwei Xiong*
92. Listening Human Behavior: 3D Human Pose Estimation With Acoustic Signals, *Yuto Shibata, Yutaka Kawashima, Mariko Isogawa, Go Irie, Akisato Kimura, Yoshimitsu Aoki*
93. Towards Domain Generalization for Multi-View 3D Object Detection in Bird-Eye-View, *Shuo Wang, Xinhai Zhao, Hai-Ming Xu, Zehui Chen, Dameng Yu, Jiahao Chang, Zhen Yang, Feng Zhao*
94. X³KD: Knowledge Distillation Across Modalities, Tasks and Stages for Multi-Camera 3D Object Detection, *Marvin Klingner, Shubhankar Borse, Varun Ravi Kumar, Behnaz Rezaei, Venkatraman Narayanan, Senthil Yogamani, Fatih Porikli*
95. Phase-Shifting Coder: Predicting Accurate Orientation in Oriented Object Detection, *Yi Yu, Feipeng Da*
96. Learned Two-Plane Perspective Prior Based Image Resampling for Efficient Object Detection, *Anurag Ghosh, N. Dinesh Reddy, Christoph Mertz, Srinivasa G. Narasimhan*
97. Resource-Efficient RGBD Aerial Tracking, *Jinyu Yang, Shang Gao, Zhe Li, Feng Zheng, Aleš Leonardis*
98. Toward RAW Object Detection: A New Benchmark and a New Model, *Ruikang Xu, Chang Chen, Jingyang Peng, Cheng Li, Yibin Huang, Fenglong Song, Youliang Yan, Zhiwei Xiong*
99. Bi-LRFusion: Bi-Directional LiDAR-Radar Fusion for 3D Dynamic Object Detection, *Yingjie Wang, Jiajun Deng, Yao Li, Jinshui Hu, Cong Liu, Yu Zhang, Jianmin Ji, Wanli Ouyang, Yanyong Zhang*
100. LiDAR-in-the-Loop Hyperparameter Optimization, *Félix Goudreault, Dominik Scheuble, Mario Bijelic, Nicolas Robidoux, Felix Heide*
101. Learning and Aggregating Lane Graphs for Urban Automated Driving, *Martin Büchner, Jannik Zörn, Ion-George Todoran, Abhinav Valada, Wolfram Burgard*
102. Center Focusing Network for Real-Time LiDAR Panoptic Segmentation, *Xiaoyan Li, Gang Zhang, Boyue Wang, Yongli Hu, Baocai Yin*
103. Adaptive Sparse Convolutional Networks With Global Context Enhancement for Faster Object Detection on Drone Images, *Bowei Du, Yecheng Huang, Jiaxin Chen, Di Huang*
104. MV-JAR: Masked Voxel Jigsaw and Reconstruction for LiDAR-Based Self-Supervised Pre-Training, *Runsen Xu, Tai Wang, Wenwei Zhang, Runjian Chen, Jinkun Cao, Jiangmiao Pang, Dahua Lin*
105. ALSO: Automotive Lidar Self-Supervision by Occupancy Estimation, *Alexandre Boulch, Corentin Sautier, Björn Michele, Gilles Puy, Renaud Marlet*
106. Unsupervised Intrinsic Image Decomposition With LiDAR Intensity, *Shogo Sato, Yasuhiro Yao, Taiga Yoshida, Takuhiro Kaneko, Shingo Ando, Jun Shimamura*
107. PVT-SSD: Single-Stage 3D Object Detector With Point-Voxel Transformer, *Honghui Yang, Wenxiao Wang, Minghao Chen, Binbin Lin, Tong He, Hua Chen, Xiaofei He, Wanli Ouyang*
108. LargeKernel3D: Scaling Up Kernels in 3D Sparse CNNs, *Yukang Chen, Jianhui Liu, Xiangyu Zhang, Xiaojuan Qi, Jiaya Jia*
109. WeatherStream: Light Transport Automation of Single Image Deweathering, *Howard Zhang, Yunhao Ba, Ethan Yang, Varan Mehra, Blake Gella, Akira Suzuki, Arnold Pfahnl, Chethan Chinder Chandrappa, Alex Wong, Achuta Kadambi*
110. Mask3D: Pre-Training 2D Vision Transformers by Learning Masked 3D Priors, *Ji Hou, Xiaoliang Dai, Zijian He, Angela Dai, Matthias Nießner*
111. DSVT: Dynamic Sparse Voxel Transformer With Rotated Sets, *Haiyang Wang, Chen Shi, Shaoshuai Shi, Meng Lei, Sen Wang, Di He, Bernt Schiele, Liwei Wang*
112. IterativePFN: True Iterative Point Cloud Filtering, *Dasith de Silva Edirimuni, Xuequan Lu, Zhiwen Shao, Gang Li, Antonio Robles-Kelly, Ying He*
113. itKD: Interchange Transfer-Based Knowledge Distillation for 3D Object Detection, *Hyeon Cho, Junyong Choi, Geonwoo Baek, Wonjun Hwang*
114. ISBNNet: A 3D Point Cloud Instance Segmentation Network With Instance-Aware Sampling and Box-Aware Dynamic Convolution, *Tuan Duc Ngo, Binh-Son Hua, Khoi Nguyen*

115. Symmetric Shape-Preserving Autoencoder for Unsupervised Real Scene Point Cloud Completion, *Changfeng Ma, Yinuo Chen, Pengxiao Guo, Jie Guo, Chongjun Wang, Yanwen Guo*
116. GeoMAE: Masked Geometric Target Prediction for Self-Supervised Point Cloud Pre-Training, *Xiaoyu Tian, Haoxi Ran, Yue Wang, Hang Zhao*
117. AnchorFormer: Point Cloud Completion From Discriminative Nodes, *Zhikai Chen, Fuchen Long, Zhaofan Qiu, Ting Yao, Wengang Zhou, Jiebo Luo, Tao Mei*
118. SHS-Net: Learning Signed Hyper Surfaces for Oriented Normal Estimation of Point Clouds, *Qing Li, Huifang Feng, Kanle Shi, Yue Gao, Yi Fang, Yu-Shen Liu, Zhizhong Han*
119. NerVE: Neural Volumetric Edges for Parametric Curve Extraction From Point Cloud, *Xiangyu Zhu, Dong Du, Weikai Chen, Zhiyou Zhao, Yinyu Nie, Xiaoguang Han*
120. Unsupervised Deep Probabilistic Approach for Partial Point Cloud Registration, *Guofeng Mei, Hao Tang, Xiaoshui Huang, Weijie Wang, Juan Liu, Jian Zhang, Luc Van Gool, Qiang Wu*
121. Local Connectivity-Based Density Estimation for Face Clustering, *Junho Shin, Hyo-Jun Lee, Hyunseop Kim, Jong-Hyeon Baek, Daehyun Kim, Yeong Jun Koh*
122. Bridging Search Region Interaction With Template for RGB-T Tracking, *Tianrui Hui, Zizheng Xun, Fengguang Peng, Junshi Huang, Xiaoming Wei, Xiaolin Wei, Jiao Dai, Jizhong Han, Si Liu*
123. Quantum Multi-Model Fitting, *Matteo Farina, Luca Magri, Willi ✧ Menapace, Elisa Ricci, Vladislav Golyanik, Federica Arrigoni*
124. Generalizable Local Feature Pre-Training for Deformable Shape ✧ Analysis, *Souhaib Attaiki, Lei Li, Maks Ovsjanikov*
125. Similarity Metric Learning for RGB-Infrared Group Re-Identification, *Jianghao Xiong, Jianhuang Lai*
126. Unsupervised Deep Asymmetric Stereo Matching With Spatially-Adaptive Self-Similarity, *Taeyong Song, Sunok Kim, Kwanghoon Sohn*
127. Sliced Optimal Partial Transport, *Yikun Bai, Bernhard Schmitzer, Matthew Thorpe, Soheil Kolouri*
128. DistractFlow: Improving Optical Flow Estimation via Realistic Distractions and Pseudo-Labeling, *Jisoo Jeong, Hong Cai, Risheek Garrepalli, Fatih Porikli*
129. Bayesian Posterior Approximation With Stochastic Ensembles, *Oleksandr Balabanov, Bernhard Mehlig, Hampus Linander*
130. V2V4Real: A Real-World Large-Scale Dataset for Vehicle-to-Vehicle Cooperative Perception, *Runsheng Xu, Xin Xia, Jinlong Li, Hanzhao Li, Shuo Zhang, Zhengzhong Tu, Zonglin Meng, Hao Xiang, Xiaoyu Dong, Rui Song, Hongkai Yu, Bolei Zhou, Jiaqi Ma*
131. ReasonNet: End-to-End Driving With Temporal and Global Reasoning, *Hao Shao, Letian Wang, Ruobing Chen, Steven L. Waslander, Hongsheng Li, Yu Liu*
132. Open-World Multi-Task Control Through Goal-Aware Representation Learning and Adaptive Horizon Prediction, *Shaofei Cai, Zihao Wang, Xiaojian Ma, Anji Liu, Yitao Liang*
133. FJMP: Factorized Joint Multi-Agent Motion Prediction Over Learned Directed Acyclic Interaction Graphs, *Luke Rowe, Martin Ethier, Eli-Henry Dykhne, Krzysztof Czarnecki*
134. Trace and Pace: Controllable Pedestrian Animation via Guided Trajectory Diffusion, *Davis Rempe, Zhengyi Luo, Xue Bin Peng, Ye Yuan, Kris Kitani, Karsten Kreis, Sanja Fidler, Or Litany*
135. Galactic: Scaling End-to-End Reinforcement Learning for Rearrangement at 100k Steps-per-Second, *Vincent-Pierre Berges, Andrew Szot, Devendra Singh Chaplot, Aaron Gokaslan, Roozbeh Mottaghi, Dhruv Batra, Eric Undersander*
136. Affordances From Human Videos as a Versatile Representation for Robotics, *Shikhar Bahl, Russell Mendonca, Lili Chen, Unnat Jain, Deepak Pathak*
137. Indiscernible Object Counting in Underwater Scenes, *Guolei Sun, Zhaochong An, Yun Liu, Ce Liu, Christos Sakaridis, Deng-Ping Fan, Luc Van Gool*
138. Tracking Through Containers and Occluders in the Wild, *Basile Van Hoorick, Pavel Tokmakov, Simon Stent, Jie Li, Carl Vondrick*
139. Simple Cues Lead to a Strong Multi-Object Tracker, *Jenny Seidenschwarz, Guillem Brasó, Víctor Castro Serrano, Ismail Elezi, Laura Leal-Taixé*
140. An In-Depth Exploration of Person Re-Identification and Gait Recognition in Cloth-Changing Conditions, *Weijia Li, Saihui Hou, Chunjie Zhang, Chunshui Cao, Xu Liu, Yongzhen Huang, Yao Zhao*
141. SelfME: Self-Supervised Motion Learning for Micro-Expression Recognition, *Xinqi Fan, Xueli Chen, Mingjie Jiang, Ali Raza Shahid, Hong Yan*
142. LipFormer: High-Fidelity and Generalizable Talking Face Generation With a Pre-Learned Facial Codebook, *Jiayu Wang, Kang Zhao, Shiwei Zhang, Yingya Zhang, Yujun Shen, Deli Zhao, Jingren Zhou*
143. Real-Time Multi-Person Eyeblick Detection in the Wild for Untrimmed Video, *Wenzheng Zeng, Yang Xiao, Sicheng Wei, Jinfang Gan, Xintao Zhang, Zhiguo Cao, Zhiwen Fang, Joey Tianyi Zhou*
144. Skinned Motion Retargeting With Residual Perception of Motion Semantics & Geometry, *Jiaxu Zhang, Junwu Weng, Di Kang, Fang Zhao, Shaoli Huang, Xuefei Zhe, Linchao Bao, Ying Shan, Jue Wang, Zhigang Tu*
145. MoDi: Unconditional Motion Synthesis From Diverse Data, *Sigal Raab, Inbal Leibovitch, Peizhuo Li, Kfir Aberman, Olga Sorkine-Hornung, Daniel Cohen-Or*
146. Recurrent Vision Transformers for Object Detection With Event Cameras, *Mathias Gehrig, Davide Scaramuzza*
147. Continuous Intermediate Token Learning With Implicit Motion Manifold for Keyframe Based Motion Interpolation, *Clinton A. Mo, Kun Hu, Chengjiang Long, Zhiyong Wang*
148. EvShutter: Transforming Events for Unconstrained Rolling Shutter Correction, *Julius Erbach, Stepan Tulyakov, Patricia Vitoria, Alfredo Bochicchio, Yuanyou Li*
149. Multi Domain Learning for Motion Magnification, *Jasdeep Singh, Subrahmanyam Murala, G. Sankara Raju Kosuru*
150. Learning Event Guided High Dynamic Range Video Reconstruction, *Yixin Yang, Jin Han, Jinxiu Liang, Imari Sato, Boxin Shi*
151. Joint Video Multi-Frame Interpolation and Deblurring Under Unknown Exposure Time, *Wei Shang, Dongwei Ren, Yi Yang, Hongzhi Zhang, Kede Ma, Wangmeng Zuo*
152. FeatER: An Efficient Network for Human Reconstruction via Feature Map-Based Transformer, *Ce Zheng, Matias Mendieta, Taojiannan Yang, Guo-Jun Qi, Chen Chen*
153. MetaFusion: Infrared and Visible Image Fusion via Meta-Feature Embedding From Object Detection, *Wenda Zhao, Shigeng Xie, Fan Zhao, You He, Huchuan Lu*

154. Joint HDR Denoising and Fusion: A Real-World Mobile HDR Image Dataset, *Shuaizheng Liu, Xindong Zhang, Lingchen Sun, Zhetong Liang, Hui Zeng, Lei Zhang*
155. Visibility Constrained Wide-Band Illumination Spectrum Design for Seeing-in-the-Dark, *Muyao Niu, Zhuoxiao Li, Zhihang Zhong, Yinqiang Zheng*
156. Self-Supervised Blind Motion Deblurring With Deep Expectation Maximization, *Ji Li, Weixi Wang, Yuesong Nan, Hui Ji*
157. Structure Aggregation for Cross-Spectral Stereo Image Guided Denoising, *Zehua Sheng, Zhu Yu, Xiongwei Liu, Si-Yuan Cao, Yuqi Liu, Hui-Liang Shen, Huaqi Zhang*
158. Rawgment: Noise-Accounted RAW Augmentation Enables Recognition in a Wide Variety of Environments, *Masakazu Yoshimura, Junji Otsuka, Atsushi Irie, Takeshi Ohashi*
159. Zero-Shot Noise2Noise: Efficient Image Denoising Without Any Data, *Youssef Mansour, Reinhard Heckel*
160. Real-Time Controllable Denoising for Image and Video, *Zhaoyang Zhang, Yitong Jiang, Wenqi Shao, Xiaogang Wang, Ping Luo, Kaimo Lin, Jinwei Gu*
161. Probability-Based Global Cross-Modal Upsampling for Pansharpening, *Zeyu Zhu, Xiangyong Cao, Man Zhou, Junhao Huang, Deyu Meng*
162. ShadowDiffusion: When Degradation Prior Meets Diffusion Model for Shadow Removal, *Lanqing Guo, Chong Wang, Wenhan Yang, Siyu Huang, Yupei Wang, Hanspeter Pfister, Bihan Wen*
163. Visual Recognition-Driven Image Restoration for Multiple Degradation With Intrinsic Semantics Recovery, *Zizheng Yang, Jie Huang, Jiahao Chang, Man Zhou, Hu Yu, Jinghao Zhang, Feng Zhao*
164. Blind Image Quality Assessment via Vision-Language Correspondence: A Multitask Learning Perspective, *Weixia Zhang, Guangtao Zhai, Ying Wei, Xiaokang Yang, Kede Ma*
165. Human Guided Ground-Truth Generation for Realistic Image Super-Resolution, *Du Chen, Jie Liang, Xindong Zhang, Ming Liu, Hui Zeng, Lei Zhang*
166. Real-Time 6K Image Rescaling With Rate-Distortion Optimization, *Chenyang Qi, Xin Yang, Ka Leong Cheng, Ying-Cong Chen, Qifeng Chen*
167. Equivalent Transformation and Dual Stream Network Construction for Mobile Image Super-Resolution, *Jiahao Chao, Zhou Zhou, Hongfan Gao, Jiali Gong, Zhengfeng Yang, Zhenbing Zeng, Lydia Dehbi*
168. Ultrahigh Resolution Image/Video Matting With Spatio-Temporal Sparsity, *Yanan Sun, Chi-Keung Tang, Yu-Wing Tai*
169. Comprehensive and Delicate: An Efficient Transformer for Image Restoration, *Haiyu Zhao, Yuanbiao Gou, Boyun Li, Dezhong Peng, Jiancheng Lv, Xi Peng*
170. PHA: Patch-Wise High-Frequency Augmentation for Transformer-Based Person Re-Identification, *Guiwei Zhang, Yongfei Zhang, Tianyu Zhang, Bo Li, Shiliang Pu*
171. PyramidFlow: High-Resolution Defect Contrastive Localization Using Pyramid Normalizing Flow, *Jiarui Lei, Xiaobo Hu, Yue Wang, Dong Liu*
172. Neural Fourier Filter Bank, *Zhijie Wu, Yuhe Jin, Kwang Moo Yi*
173. Restoration of Hand-Drawn Architectural Drawings Using Latent Space Mapping With Degradation Generator, *Nakkwan Choi, Seungjae Lee, Yongsik Lee, Seungjoon Yang*
174. Neural Preset for Color Style Transfer, *Zhanghan Ke, Yuhao Liu, Lei Zhu, Nanxuan Zhao, Rynson W.H. Lau*
175. NÜWA-LIP: Language-Guided Image Inpainting With Defect-Free VQGAN, *Minheng Ni, Xiaoming Li, Wangmeng Zuo*
176. DualVector: Unsupervised Vector Font Synthesis With Dual-Part Representation, *Ying-Tian Liu, Zhifei Zhang, Yuan-Chen Guo, Matthew Fisher, Zhaowen Wang, Song-Hai Zhang*
177. DATID-3D: Diversity-Preserved Domain Adaptation Using Text-to-Image Diffusion for 3D Generative Model, *Gwanghyun Kim, Se Young Chun*
178. GALIP: Generative Adversarial CLIPs for Text-to-Image Synthesis, *Ming Tao, Bing-Kun Bao, Hao Tang, Changsheng Xu*
179. Fix the Noise: Disentangling Source Feature for Controllable Domain Translation, *Dongyeun Lee, Jae Young Lee, Doyeon Kim, Jaehyun Choi, Jaeyun Yoo, Junmo Kim*
180. Conditional Text Image Generation With Diffusion Models, *Yuanzhi Zhu, Zhaohai Li, Tianwei Wang, Mengchao He, Cong Yao*
181. ReCo: Region-Controlled Text-to-Image Generation, *Zhengyuan Yang, Jianfeng Wang, Zhe Gan, Linjie Li, Kevin Lin, Chenfei Wu, Nan Duan, Zicheng Liu, Ce Liu, Michael Zeng, Lijuan Wang*
182. Freestyle Layout-to-Image Synthesis, *Han Xue, Zhiwu Huang, ✧ Qianru Sun, Li Song, Wenjun Zhang*
183. Specialist Diffusion: Plug-and-Play Sample-Efficient Fine-Tuning of Text-to-Image Diffusion Models to Learn Any Unseen Style, *Haoming Lu, Hazarapet Tunanyan, Kai Wang, Shant Navasardyan, Zhangyang Wang, Humphrey Shi*
184. Toward Verifiable and Reproducible Human Evaluation for Text-to-Image Generation, *Mayu Otani, Riku Togashi, Yu Sawai, Ryosuke Ishigami, Yuta Nakashima, Esa Rahtu, Janne Heikkilä, Shin'ichi Satoh*
185. Towards Flexible Multi-Modal Document Models, *Naoto Inoue, ✧ Kotaro Kikuchi, Edgar Simo-Serra, Mayu Otani, Kota Yamaguchi*
186. On Distillation of Guided Diffusion Models, *Chenlin Meng, Robin Rombach, Ruiqi Gao, Diederik Kingma, Stefano Ermon, Jonathan Ho, Tim Salimans*
187. Dimensionality-Varying Diffusion Process, *Han Zhang, Ruili Feng, Zhantao Yang, Lianghua Huang, Yu Liu, Yifei Zhang, Yujun Shen, Deli Zhao, Jingren Zhou, Fan Cheng*
188. Shape-Aware Text-Driven Layered Video Editing, *Yao-Chih Lee, Ji-Ze Genevieve Jang, Yi-Ting Chen, Elizabeth Qiu, Jia-Bin Huang*
189. Rethinking Image Super Resolution From Long-Tailed Distribution Learning Perspective, *Yuanbiao Gou, Peng Hu, Jiancheng Lv, Hongyuan Zhu, Xi Peng*
190. End-to-End Video Matting With Trimap Propagation, *Wei-Lun Huang, Ming-Sui Lee*
191. Context-Based Trit-Plane Coding for Progressive Image Compression, *Seungmin Jeon, Kwang Pyo Choi, Youngo Park, Chang-Su Kim*
192. Complexity-Guided Slimmable Decoder for Efficient Deep Video Compression, *Zhihao Hu, Dong Xu*
193. Efficient Hierarchical Entropy Model for Learned Point Cloud Compression, *Rui Song, Chunyang Fu, Shan Liu, Ge Li*
194. NIRVANA: Neural Implicit Representations of Videos With Adaptive Networks and Autoregressive Patch-Wise Modeling, *Shishira R. Maiya, Sharath Girish, Max Ehrlich, Hanyu Wang, Kwot Sin Lee, Patrick Poirson, Pengxiang Wu, Chen Wang, Abhinav Shrivastava*
195. Learned Image Compression With Mixed Transformer-CNN Architectures, *Jinming Liu, Heming Sun, Jiro Katto*

196. Memory-Friendly Scalable Super-Resolution via Rewinding Lottery Ticket Hypothesis, *Jin Lin, Xiaotong Luo, Ming Hong, Yanyun Qu, Yuan Xie, Zongze Wu*
197. InternImage: Exploring Large-Scale Vision Foundation Models ✧ With Deformable Convolutions, *Wenhai Wang, Jifeng Dai, Zhe Chen, Zhenhang Huang, Zhiqi Li, Xizhou Zhu, Xiaowei Hu, Tong Lu, Lewei Lu, Hongsheng Li, Xiaogang Wang, Yu Qiao*
198. EfficientViT: Memory Efficient Vision Transformer With Cascaded Group Attention, *Xinyu Liu, Houwen Peng, Ningxin Zheng, Yuqing Yang, Han Hu, Yixuan Yuan*
199. Castling-ViT: Compressing Self-Attention via Switching Towards Linear-Angular Attention at Vision Transformer Inference, *Haoran You, Yunsang Xiong, Xiaoliang Dai, Bichen Wu, Peizhao Zhang, Haoqi Fan, Peter Vajda, Yingyan (Celine) Lin*
200. RIFormer: Keep Your Vision Backbone Effective but Removing Token Mixer, *Jiahao Wang, Songyang Zhang, Yong Liu, Taiqiang Wu, Yujiu Yang, Xihui Liu, Kai Chen, Ping Luo, Dahua Lin*
201. High-Resolution Image Reconstruction With Latent Diffusion Models From Human Brain Activity, *Yu Takagi, Shinji Nishimoto*
202. Non-Contrastive Unsupervised Learning of Physiological ✧ Signals From Video, *Jeremy Speth, Nathan Vance, Patrick Flynn, Adam Czajka*
203. Revealing the Dark Secrets of Masked Image Modeling, *Zhenda Xie, Zigang Geng, Jingcheng Hu, Zheng Zhang, Han Hu, Yue Cao*
204. Improving Visual Representation Learning Through Perceptual Understanding, *Samyakh Tukra, Frederick Hoffman, Ken Chatfield*
205. FlexiViT: One Model for All Patch Sizes, *Lucas Beyer, Pavel Izmailov, Alexander Kolesnikov, Mathilde Caron, Simon Kornblith, Xiaohua Zhai, Matthias Minderer, Michael Tschannen, Ibrahim Alabdulmohsin, Filip Pavetic*
206. AdaMAE: Adaptive Masking for Efficient Spatiotemporal Learning With Masked Autoencoders, *Wele Gedara Chaminda Bandara, Naman Patel, Ali Gholami, Mehdi Nikkhah, Motilal Agrawal, Vishal M. Patel*
207. SimpSON: Simplifying Photo Cleanup With Single-Click Distracting Object Segmentation Network, *Chuonq Huynh, Yuqian Zhou, Zhe Lin, Connelly Barnes, Eli Shechtman, Sohrab Amirghodsi, Abhinav Shrivastava*
208. Visual Dependency Transformers: Dependency Tree Emerges From Reversed Attention, *Mingyu Ding, Yikang Shen, Lijie Fan, Zhenfang Chen, Zitian Chen, Ping Luo, Joshua B. Tenenbaum, Chuang Gan*
209. Iterative Next Boundary Detection for Instance Segmentation of Tree Rings in Microscopy Images of Shrub Cross Sections, *Alexander Gillert, Giulia Resente, Alba Anadon-Rosell, Martin Wilmking, Uwe Freiherr von Lukas*
210. VideoMAE V2: Scaling Video Masked Autoencoders With Dual Masking, *Limin Wang, Bingkun Huang, Zhiyu Zhao, Zhan Tong, Yinan He, Yi Wang, Yali Wang, Yu Qiao*
211. DropMAE: Masked Autoencoders With Spatial-Attention Dropout for Tracking Tasks, *Qiangqiang Wu, Tianyu Yang, Ziquan Liu, Baoyuan Wu, Ying Shan, Antoni B. Chan*
212. SeqTrack: Sequence to Sequence Learning for Visual Object Tracking, *Xin Chen, Houwen Peng, Dong Wang, Huchuan Lu, Han Hu*
213. Bootstrapping Objectness From Videos by Relaxed Common Fate and Visual Grouping, *Long Lian, Zhirong Wu, Stella X. Yu*
214. Video Event Restoration Based on Keyframes for Video Anomaly Detection, *Zhiwei Yang, Jing Liu, Zhaoyang Wu, Peng Wu, Xiaotao Liu*
215. Streaming Video Model, *Yucheng Zhao, Chong Luo, Chuanxin Tang, Dongdong Chen, Noel Codella, Zheng-Jun Zha*
216. LSTFE-Net: Long Short-Term Feature Enhancement Network for Video Small Object Detection, *Jinsheng Xiao, Yuanxu Wu, Yunhua Chen, Shurui Wang, Zhongyuan Wang, Jiayi Ma*
217. A Generalized Framework for Video Instance Segmentation, *Miran Heo, Sukjun Hwang, Jeongseok Hyun, Hanjung Kim, Seoung Wug Oh, Joon-Young Lee, Seon Joo Kim*
218. Referring Multi-Object Tracking, *Dongming Wu, Wencheng Han, Tiancai Wang, Xingping Dong, Xiangyu Zhang, Jianbing Shen*
219. Source-Free Video Domain Adaptation With Spatial-Temporal-Historical Consistency Learning, *Kai Li, Deep Patel, Erik Kruus, Martin Renqiang Min*
220. Seeing What You Said: Talking Face Generation Guided by a Lip Reading Expert, *Jiadong Wang, Xinyuan Qian, Malu Zhang, Robby T. Tan, Haizhou Li*
221. Egocentric Auditory Attention Localization in Conversations, *Fiona Ryan, Hao Jiang, Abhinav Shukla, James M. Rehg, Vamsi Krishna Ithapu*
222. iQuery: Instruments As Queries for Audio-Visual Sound Separation, *Jiabao Chen, Renrui Zhang, Dongze Lian, Jiaqi Yang, Ziyao Zeng, Jianbo Shi*
223. Learning to Dub Movies via Hierarchical Prosody Models, *Gaoxiang Cong, Liang Li, Yuankai Qi, Zheng-Jun Zha, Qi Wu, Wenyu Wang, Bin Jiang, Ming-Hsuan Yang, Qingming Huang*
224. A Large-Scale Robustness Analysis of Video Action Recognition Models, *Madeline Chantry Schiappa, Naman Biyani, Prudvi Kamtam, Shruti Vyas, Hamid Palangi, Vibhav Vineet, Yogesh S. Rawat*
225. The Wisdom of Crowds: Temporal Progressive Attention for Early Action Prediction, *Alexandros Stergiou, Dima Damen*
226. STMixer: A One-Stage Sparse Action Detector, *Tao Wu, Mengqi Cao, Ziteng Gao, Gangshan Wu, Limin Wang*
227. Generating Human Motion From Textual Descriptions With Discrete Representations, *Jianrong Zhang, Yangsong Zhang, Xiaodong Cun, Yong Zhang, Hongwei Zhao, Hongtao Lu, Xi Shen, Ying Shan*
228. Cascade Evidential Learning for Open-World Weakly-Supervised Temporal Action Localization, *Mengyuan Chen, Junyu Gao, Changsheng Xu*
229. Distilling Vision-Language Pre-Training to Collaborate With Weakly-Supervised Temporal Action Localization, *Chen Ju, Kunhao Zheng, Jinxiang Liu, Peisen Zhao, Ya Zhang, Jianlong Chang, Qi Tian, Yanfeng Wang*
230. Simultaneously Short- and Long-Term Temporal Modeling for Semi-Supervised Video Semantic Segmentation, *Jiangwei Lao, Weixiang Hong, Xin Guo, Yingying Zhang, Jian Wang, Jingdong Chen, Wei Chu*
231. MIST: Multi-Modal Iterative Spatial-Temporal Transformer for Long-Form Video Question Answering, *Difei Gao, Luowei Zhou, Lei Ji, Linchao Zhu, Yi Yang, Mike Zheng Shou*
232. ✧ Language-Guided Music Recommendation for Video via Prompt Analogies, *Daniel McKee, Justin Salamon, Josef Sivic, Bryan Russell*

233. Text-Visual Prompting for Efficient 2D Temporal Video Grounding, *Yimeng Zhang, Xin Chen, Jinghan Jia, Sijia Liu, Ke Ding*
234. CelebV-Text: A Large-Scale Facial Text-Video Dataset, *Jianhui Yu, Hao Zhu, Liming Jiang, Chen Change Loy, Weidong Cai, Wayne Wu*
235. CNVid-3.5M: Build, Filter, and Pre-Train the Large-Scale Public Chinese Video-Text Dataset, *Tian Gan, Qing Wang, Xingning Dong, Xiangyuan Ren, Liqiang Nie, Qingpei Guo*
236. Learning Procedure-Aware Video Representation From Instructional Videos and Their Narrations, *Yiwu Zhong, Licheng Yu, Yang Bai, Shangwen Li, Xueting Yan, Yin Li*
237. PDPP: Projected Diffusion for Procedure Planning in ✨ Instructional Videos, *Hanlin Wang, Yilu Wu, Sheng Guo, Limin Wang*
238. Towards Fast Adaptation of Pretrained Contrastive Models for Multi-Channel Video-Language Retrieval, *Xudong Lin, Simran Tiwari, Shiyuan Huang, Manling Li, Mike Zheng Shou, Heng Ji, Shih-Fu Chang*
239. Clover: Towards a Unified Video-Language Alignment and Fusion Model, *Jingjia Huang, Yinan Li, Jiashi Feng, Xinglong Wu, Xiaoshuai Sun, Rongrong Ji*
240. Align and Attend: Multimodal Summarization With Dual Contrastive Losses, *Bo He, Jun Wang, Jielin Qiu, Trung Bui, Abhinav Shrivastava, Zhaowen Wang*
241. Learning Situation Hyper-Graphs for Video Question Answering, *Aisha Urooj, Hilde Kuehne, Bo Wu, Kim Chheu, Walid Bousselham, Chuang Gan, Niels Lobo, Mubarak Shah*
242. Natural Language-Assisted Sign Language Recognition, *Ronglai Zuo, Fangyun Wei, Brian Mak*
243. SkyEye: Self-Supervised Bird's-Eye-View Semantic Mapping Using Monocular Frontal View Images, *Nikhil Gosala, Kürsat Petek, Paulo L. J. Drews-Jr, Wolfram Burgard, Abhinav Valada*
244. Adaptive Zone-Aware Hierarchical Planner for Vision-Language Navigation, *Chen Gao, Xingyu Peng, Mi Yan, He Wang, Lirong Yang, Haibing Ren, Hongsheng Li, Si Liu*
245. Iterative Vision-and-Language Navigation, *Jacob Krantz, Shurjo Banerjee, Wang Zhu, Jason Corso, Peter Anderson, Stefan Lee, Jesse Thomason*
246. EXCALIBUR: Encouraging and Evaluating Embodied Exploration, *Hao Zhu, Raghav Kapoor, So Yeon Min, Winson Han, Jiatai Li, Kaiwen Geng, Graham Neubig, Yonatan Bisk, Aniruddha Kembhavi, Luca Weihs*
247. Multimodal Prompting With Missing Modalities for Visual Recognition, *Yi-Lun Lee, Yi-Hsuan Tsai, Wei-Chen Chiu, Chen-Yu Lee*
248. Visual Programming: Compositional Visual Reasoning Without 🌀 Training, *Tanmay Gupta, Aniruddha Kembhavi*
249. Super-CLEVR: A Virtual Benchmark to Diagnose Domain ✨ Robustness in Visual Reasoning, *Zhuowan Li, Xingrui Wang, Elias Stengel-Eskin, Adam Kortylewski, Wufei Ma, Benjamin Van Durme, Alan L. Yuille*
250. Prompting Large Language Models With Answer Heuristics for Knowledge-Based Visual Question Answering, *Zhenwei Shao, Zhou Yu, Meng Wang, Jun Yu*
251. À-La-Carte Prompt Tuning (APT): Combining Distinct Data via Composable Prompting, *Benjamin Bowman, Alessandro Achille, Luca Zancato, Matthew Trager, Pramuditha Perera, Giovanni Paolini, Stefano Soatto*
252. ConStruct-VL: Data-Free Continual Structured VL Concepts Learning, *James Seale Smith, Paola Cascante-Bonilla, Assaf Arbelle, Donghyun Kim, Rameswar Panda, David Cox, Diyi Yang, Zsolt Kira, Rogerio Feris, Leonid Karlinsky*
253. Q: How to Specialize Large Vision-Language Models to Data-Scarce VQA Tasks? A: Self-Train on Unlabeled Images!, *Zaid Khan, Vijay Kumar BG, Samuel Schuster, Xiang Yu, Yun Fu, Manmohan Chandraker*
254. Learning to Exploit Temporal Structure for Biomedical Vision-Language Processing, *Shruthi Bannur, Stephanie Hyland, Qianchu Liu, Fernando Pérez-García, Maximilian Ilse, Daniel C. Castro, Benedikt Boecking, Harshita Sharma, Kenza Bouzid, Anja Thieme, Anton Schwaighofer, Maria Wetscherek, Matthew P. Lungren, Aditya Nori, Javier Alvarez-Valle, Ozan Oktay*
255. FashionSAP: Symbols and Attributes Prompt for Fine-Grained Fashion Vision-Language Pre-Training, *Yunpeng Han, Lisai Zhang, Qingcai Chen, Zhijian Chen, Zhonghua Li, Jianxin Yang, Zhao Cao*
256. Advancing Visual Grounding With Scene Knowledge: Benchmark and Method, *Zhihong Chen, Ruifei Zhang, Yibing Song, Xiang Wan, Guanbin Li*
257. Beyond Appearance: A Semantic Controllable Self-Supervised Learning Framework for Human-Centric Visual Tasks, *Weihua Chen, Xianzhe Xu, Jian Jia, Hao Luo, Yaohua Wang, Fan Wang, Rong Jin, Xiuyu Sun*
258. OCTET: Object-Aware Counterfactual Explanations, *Mehdi Zemni, Mickaël Chen, Éloi Zablocki, Hédi Ben-Younes, Patrick Pérez, Matthieu Cord*
259. Local-Guided Global: Paired Similarity Representation for Visual Reinforcement Learning, *Hyesong Choi, Hunsang Lee, Wonil Song, Sangryul Jeon, Kwanghoon Sohn, Dongbo Min*
260. What Can Human Sketches Do for Object Detection? *Pinaki 🌀 Nath Chowdhury, Ayan Kumar Bhunia, Aneeshan Sain, Subhadeep Koley, Tao Xiang, Yi-Zhe Song*
261. Revisiting Multimodal Representation in Contrastive Learning: From Patch and Token Embeddings to Finite Discrete Tokens, *Yuxiao Chen, Jianbo Yuan, Yu Tian, Shijie Geng, Xinyu Li, Ding Zhou, Dimitris N. Metaxas, Hongxia Yang*
262. Correlational Image Modeling for Self-Supervised Visual Pre-Training, *Wei Li, Jiahao Xie, Chen Change Loy*
263. Generalized Decoding for Pixel, Image, and Language, *Xueyan Zou, Zi-Yi Dou, Jianwei Yang, Zhe Gan, Linjie Li, Chunyuan Li, Xiyang Dai, Harkirat Behl, Jianfeng Wang, Lu Yuan, Nanyun Peng, Lijuan Wang, Yong Jae Lee, Jianfeng Gao*
264. Towards Modality-Agnostic Person Re-Identification With Descriptive Query, *Cuiqun Chen, Mang Ye, Ding Jiang*
265. M⁶Doc: A Large-Scale Multi-Format, Multi-Type, Multi-Layout, Multi-Language, Multi-Annotation Category Dataset for Modern Document Layout Analysis, *Hiuyi Cheng, Peirong Zhang, Sihang Wu, Jiaxin Zhang, Qiyuan Zhu, Zecheng Xie, Jing Li, Kai Ding, Lianwen Jin*
266. Learning Customized Visual Models With Retrieval-Augmented ✨ Knowledge, *Haotian Liu, Kilho Son, Jianwei Yang, Ce Liu, Jianfeng Gao, Yong Jae Lee, Chunyuan Li*
267. Learning Semantic Relationship Among Instances for Image-Text Matching, *Zheren Fu, Zhendong Mao, Yan Song, Yongdong Zhang*

268. l2MVFormer: Large Language Model Generated Multi-View Document Supervision for Zero-Shot Image Classification, *Muhammad Ferjad Naeem, Muhammad Gul Zain Ali Khan, Yongqin Xian, Muhammad Zeshan Afzal, Didier Stricker, Luc Van Gool, Federico Tombari*
269. ImageBind: One Embedding Space to Bind Them All, *Rohit Girdhar, Alaaeldin El-Nouby, Zhuang Liu, Mannat Singh, Kalyan Vasudev Alwala, Armand Joulin, Ishan Misra*
270. Model-Agnostic Gender Debaised Image Captioning, *Yusuke Hirota, Yuta Nakashima, Noa Garcia*
271. Boundary-Aware Backward-Compatible Representation via Adversarial Learning in Image Retrieval, *Tan Pan, Furong Xu, Xudong Yang, Sifeng He, Chen Jiang, Qingpei Guo, Feng Qian, Xiaobo Zhang, Yuan Cheng, Lei Yang, Wei Chu*
272. Prompt, Generate, Then Cache: Cascade of Foundation Models Makes Strong Few-Shot Learners, *Renrui Zhang, Xiangfei Hu, Bohao Li, Siyuan Huang, Hanqiu Deng, Yu Qiao, Peng Gao, Hongsheng Li*
273. Towards Unified Scene Text Spotting Based on Sequence Generation, *Taeho Kil, Seonghyeon Kim, Sukmin Seo, Yoonsik Kim, Daehee Kim*
274. CapDet: Unifying Dense Captioning and Open-World Detection Pretraining, *Yanxin Long, Youpeng Wen, Jianhua Han, Hang Xu, Pengzhen Ren, Wei Zhang, Shen Zhao, Xiaodan Liang*
275. CLIP²: Contrastive Language-Image-Point Pretraining From Real-World Point Cloud Data, *Yihan Zeng, Chenhan Jiang, Jiageng Mao, Jianhua Han, Chaoqiang Ye, Qingqiu Huang, Dit-Yan Yeung, Zhen Yang, Xiaodan Liang, Hang Xu*
276. Aligning Bag of Regions for Open-Vocabulary Object Detection, *Size Wu, Wenwei Zhang, Sheng Jin, Wentao Liu, Chen Change Loy*
277. Visual Recognition by Request, *Chufeng Tang, Lingxi Xie, Xiaopeng Zhang, Xiaolin Hu, Qi Tian*
278. Category Query Learning for Human-Object Interaction Classification, *Chi Xie, Fangao Zeng, Yue Hu, Shuang Liang, Yichen Wei*
279. Self-Supervised Implicit Glyph Attention for Text Recognition, *Tongkun Guan, Chaochen Gu, Jingzheng Tu, Xue Yang, Qi Feng, Yudi Zhao, Wei Shen*
280. Enlarging Instance-Specific and Class-Specific Information for Open-Set Action Recognition, *Jun Cen, Shiwei Zhang, Xiang Wang, Yixuan Pei, Zhiwu Qing, Yingya Zhang, Qifeng Chen*
281. CLIP Is Also an Efficient Segmenter: A Text-Driven Approach for Weakly Supervised Semantic Segmentation, *Yuqi Lin, Minghao Chen, Wenxiao Wang, Boxi Wu, Ke Li, Binbin Lin, Haifeng Liu, Xiaofei He*
282. Learning Attention As Disentangler for Compositional Zero-Shot Learning, *Shaozhe Hao, Kai Han, Kwan-Yee K. Wong*
283. Universal Instance Perception As Object Discovery and Retrieval, *Bin Yan, Yi Jiang, Jiannan Wu, Dong Wang, Ping Luo, Zehuan Yuan, Huchuan Lu*
284. Progressive Semantic-Visual Mutual Adaption for Generalized Zero-Shot Learning, *Man Liu, Feng Li, Chunjie Zhang, Yunchao Wei, Huihui Bai, Yao Zhao*
285. DPF: Learning Dense Prediction Fields With Weak Supervision, *Xiaoxue Chen, Yuhang Zheng, Yupeng Zheng, Qiang Zhou, Hao Zhao, Guyue Zhou, Ya-Qin Zhang*
286. Modeling Entities As Semantic Points for Visual Information Extraction in the Wild, *Zhibo Yang, Rujiao Long, Pengfei Wang, Sibao Song, Humen Zhong, Wenqing Cheng, Xiang Bai, Cong Yao*
287. GeoNet: Benchmarking Unsupervised Adaptation Across Geographies, *Tarun Kalluri, Wangdong Xu, Manmohan Chandraker*
288. SegLoc: Learning Segmentation-Based Representations for Privacy-Preserving Visual Localization, *Maxime Pietrantonni, Martin Humenberger, Torsten Sattler, Gabriela Csurka*
289. Towards Open-World Segmentation of Parts, *Tai-Yu Pan, Qing Liu, Wei-Lun Chao, Brian Price*
290. Pruning Parameterization With Bi-Level Optimization for Efficient Semantic Segmentation on the Edge, *Changdi Yang, Pu Zhao, Yanyu Li, Wei Niu, Jiexiong Guan, Hao Tang, Minghai Qin, Bin Ren, Xue Lin, Yanzhi Wang*
291. HGFormer: Hierarchical Grouping Transformer for Domain Generalized Semantic Segmentation, *Jian Ding, Nan Xue, Gui-Song Xia, Bernt Schiele, Dengxin Dai*
292. Exemplar-FreeSOLO: Enhancing Unsupervised Instance Segmentation With Exemplars, *Taooseef Ishtiaq, Qing En, Yuhong Guo*
293. Weakly-Supervised Domain Adaptive Semantic Segmentation With Prototypical Contrastive Learning, *Anurag Das, Yongqin Xian, Dengxin Dai, Bernt Schiele*
294. Spatial-Temporal Concept Based Explanation of 3D ConvNets, *Ying Ji, Yu Wang, Jien Kato*
295. Sparsely Annotated Semantic Segmentation With Adaptive Gaussian Mixtures, *Linshan Wu, Zhun Zhong, Leyuan Fang, Xingxin He, Qiang Liu, Jiayi Ma, Hao Chen*
296. Fuzzy Positive Learning for Semi-Supervised Semantic Segmentation, *Pengchong Qiao, Zhidan Wei, Yu Wang, Zhennan Wang, Guoli Song, Fan Xu, Xiangyang Ji, Chang Liu, Jie Chen*
297. STAR Loss: Reducing Semantic Ambiguity in Facial Landmark Detection, *Zhenglin Zhou, Huaxia Li, Hong Liu, Nanyang Wang, Gang Yu, Rongrong Ji*
298. Boosting Low-Data Instance Segmentation by Unsupervised Pre-Training With Saliency Prompt, *Hao Li, Dingwen Zhang, Nian Liu, Lechao Cheng, Yalun Dai, Chao Zhang, Xinggong Wang, Junwei Han*
299. Decoupled Semantic Prototypes Enable Learning From Diverse Annotation Types for Semi-Weakly Segmentation in Expert-Driven Domains, *Simon Reiß, Constantin Seibold, Alexander Freytag, Erik Rodner, Rainer Stiefelhagen*
300. The Treasure Beneath Multiple Annotations: An Uncertainty-Aware Edge Detector, *Caixia Zhou, Yaping Huang, Mengyang Pu, Qingji Guan, Li Huang, Haibin Ling*
301. Knowledge Combination to Learn Rotated Detection Without Rotated Annotation, *Tianyu Zhu, Bryce Ferenczi, Pulak Purkait, Tom Drummond, Hamid Rezaatofghi, Anton van den Hengel*
302. Mapping Degeneration Meets Label Evolution: Learning Infrared Small Target Detection With Single Point Supervision, *Xinyi Ying, Li Liu, Yingqian Wang, Ruojing Li, Nuo Chen, Zaiping Lin, Weidong Sheng, Shilin Zhou*
303. SAP-DETR: Bridging the Gap Between Salient Points and Queries-Based Transformer Detector for Fast Model Convergence, *Yang Liu, Yao Zhang, Yixin Wang, Yang Zhang, Jiang Tian, Zhongchao Shi, Jianping Fan, Zhiqiang He*
304. Zero-Shot Object Counting, *Jingyi Xu, Hieu Le, Vu Nguyen, Viresh Ranjan, Dimitris Samaras*

305. SOOD: Towards Semi-Supervised Oriented Object Detection, *Wei Hua, Dingkan Liang, Jingyu Li, Xiaolong Liu, Zhikang Zou, Xiaoqing Ye, Xiang Bai*
306. Large-Scale Training Data Search for Object Re-Identification, *Yue Yao, Tom Gedeon, Liang Zheng*
307. Ambiguity-Resistant Semi-Supervised Learning for Dense Object Detection, *Chang Liu, Weiming Zhang, Xiangru Lin, Wei Zhang, Xiao Tan, Junyu Han, Xiaomao Li, Errui Ding, Jingdong Wang*
308. Towards Effective Visual Representations for Partial-Label Learning, *Shiyu Xia, Jiaqi Lv, Ning Xu, Gang Niu, Xin Geng*
309. Bi3D: Bi-Domain Active Learning for Cross-Domain 3D Object Detection, *Jiakang Yuan, Bo Zhang, Xiangchao Yan, Tao Chen, Botian Shi, Yikang Li, Yu Qiao*
310. Boosting Detection in Crowd Analysis via Underutilized Output Features, *Shaokai Wu, Fengyu Yang*
311. Self-Supervised Learning From Images With a Joint-Embedding Predictive Architecture, *Mahmoud Assran, Quentin Duval, Ishan Misra, Piotr Bojanowski, Pascal Vincent, Michael Rabbat, Yann LeCun, Nicolas Ballas*
312. Weakly Supervised Segmentation With Point Annotations for Histopathology Images via Contrast-Based Variational Model, *Hongrun Zhang, Liam Burrows, Yanda Meng, Declan Sculthorpe, Abhik Mukherjee, Sarah E. Coupland, Ke Chen, Yalin Zheng*
313. DoNet: Deep De-Overlapping Network for Cytology Instance Segmentation, *Hao Jiang, Rushan Zhang, Yanning Zhou, Yumeng Wang, Hao Chen*
314. MCF: Mutual Correction Framework for Semi-Supervised Medical Image Segmentation, *Yongchao Wang, Bin Xiao, Xiuli Bi, Weisheng Li, Xinbo Gao*
315. Histopathology Whole Slide Image Analysis With Heterogeneous Graph Representation Learning, *Tsai Hor Chan, Fernando Julio Cendra, Lan Ma, Guosheng Yin, Lequan Yu*
316. PEFAT: Boosting Semi-Supervised Medical Image Classification via Pseudo-Loss Estimation and Feature Adversarial Training, *Qingjie Zeng, Yutong Xie, Zilin Lu, Yong Xia*
317. Causally-Aware Intraoperative Imputation for Overall Survival Time Prediction, *Xiang Li, Xuelin Qian, Litian Liang, Lingjie Kong, Qiaole Dong, Jiejun Chen, Dingxia Liu, Xiuzhong Yao, Yanwei Fu*
318. Balanced Energy Regularization Loss for Out-of-Distribution Detection, *Hyunjun Choi, Hawook Jeong, Jin Young Choi*
319. Block Selection Method for Using Feature Norm in Out-of-Distribution Detection, *Yeonguk Yu, Sungho Shin, Seongju Lee, Changhyun Jun, Kyoobin Lee*
320. Highly Confident Local Structure Based Consensus Graph Learning for Incomplete Multi-View Clustering, *Jie Wen, Chengliang Liu, Gehui Xu, Zhihao Wu, Chao Huang, Lunke Fei, Yong Xu*
321. Siamese DETR, *Zeren Chen, Gengshi Huang, Wei Li, Jianing Teng, Kun Wang, Jing Shao, Chen Change Loy, Lu Sheng*
322. Towards Bridging the Performance Gaps of Joint Energy-Based Models, *Xiulong Yang, Qing Su, Shihao Ji*
323. Three Guidelines You Should Know for Universally Slimmable Self-Supervised Learning, *Yun-Hao Cao, Peiqin Sun, Shuchang Zhou*
324. Boosting Transductive Few-Shot Fine-Tuning With Margin-Based Uncertainty Weighting and Probability Regularization, *Ran Tao, Hao Chen, Marios Savvides*
325. CHMATCH: Contrastive Hierarchical Matching and Robust Adaptive Threshold Boosted Semi-Supervised Learning, *Jianlong Wu, Haozhe Yang, Tian Gan, Ning Ding, Feijun Jiang, Liqiang Nie*
326. MarginMatch: Improving Semi-Supervised Learning with Pseudo-Margins, *Tiberiu Sosea, Cornelia Caragea*
327. Ranking Regularization for Critical Rare Classes: Minimizing False Positives at a High True Positive Rate, *Kiarash Mohammadi, He Zhao, Mengyao Zhai, Frederick Tung*
328. Learning Imbalanced Data With Vision Transformers, *Zhengzhuo Xu, Ruikang Liu, Shuo Yang, Zenghao Chai, Chun Yuan*
329. No One Left Behind: Improving the Worst Categories in Long-Tailed Learning, *Yingxiao Du, Jianxin Wu*
330. Global and Local Mixture Consistency Cumulative Learning for Long-Tailed Visual Recognitions, *Fei Du, Peng Yang, Qi Jia, Fengtao Nan, Xiaoting Chen, Yun Yang*
331. Curvature-Balanced Feature Manifold Learning for Long-Tailed Classification, *Yanbiao Ma, Licheng Jiao, Fang Liu, Shuyuan Yang, Xu Liu, Lingling Li*
332. DAA: A Delta Age AdaIN Operation for Age Estimation via Binary Code Transformer, *Ping Chen, Xingpeng Zhang, Ye Li, Ju Tao, Bin Xiao, Bing Wang, Zongjie Jiang*
333. DLBD: A Self-Supervised Direct-Learned Binary Descriptor, *Bin Xiao, Yang Hu, Bo Liu, Xiuli Bi, Weisheng Li, Xinbo Gao*
334. Progressive Open Space Expansion for Open-Set Model Attribution, *Tianyun Yang, Danding Wang, Fan Tang, Xinying Zhao, Juan Cao, Sheng Tang*
335. DiGA: Distil to Generalize and Then Adapt for Domain Adaptive Semantic Segmentation, *Fengyi Shen, Akhil Gurram, Ziyuan Liu, He Wang, Alois Knoll*
336. Multi-Modal Learning With Missing Modality via Shared-Specific Feature Modelling, *Hu Wang, Yuanhong Chen, Congbo Ma, Jodie Avery, Louise Hull, Gustavo Carneiro*
337. Towards All-in-One Pre-Training via Maximizing Multi-Modal Mutual Information, *Weijie Su, Xizhou Zhu, Chenxin Tao, Lewei Lu, Bin Li, Gao Huang, Yu Qiao, Xiaogang Wang, Jie Zhou, Jifeng Dai*
338. Bi-Level Meta-Learning for Few-Shot Domain Generalization, *Xiaorong Qin, Xinhang Song, Shuqiang Jiang*
339. Train/Test-Time Adaptation With Retrieval, *Luca Zancato, Alessandro Achille, Tian Yu Liu, Matthew Trager, Pramuditha Perera, Stefano Soatto*
340. Robust Test-Time Adaptation in Dynamic Scenarios, *Longhui Yuan, Binhui Xie, Shuang Li*
341. Domain Expansion of Image Generators, *Yotam Nitzan, Michaël Gharbi, Richard Zhang, Taesung Park, Jun-Yan Zhu, Daniel Cohen-Or, Eli Shechtman*
342. Switchable Representation Learning Framework With Self-Compatibility, *Shengsen Wu, Yan Bai, Yihang Lou, Xiongkun Linghu, Jianzhong He, Ling-Yu Duan*
343. A New Benchmark: On the Utility of Synthetic Data With Blender for Bare Supervised Learning and Downstream Domain Adaptation, *Hui Tang, Kui Jia*
344. Adapting Shortcut With Normalizing Flow: An Efficient Tuning Framework for Visual Recognition, *Yaoming Wang, Bowen Shi, Xiaopeng Zhang, Jin Li, Yuchen Liu, Wenrui Dai, Chenglin Li, Hongkai Xiong, Qi Tian*

345. Manipulating Transfer Learning for Property Inference, *Yulong Tian, Fnu Suya, Anshuman Suri, Fengyuan Xu, David Evans*
346. Heterogeneous Continual Learning, *Divyam Madaan, Hongxu Yin, Wonmin Byeon, Jan Kautz, Pavlo Molchanov*
347. Generic-to-Specific Distillation of Masked Autoencoders, *Wei Huang, Zhiliang Peng, Li Dong, Furu Wei, Jianbin Jiao, Qixiang Ye*
348. Towards a Smaller Student: Capacity Dynamic Distillation for Efficient Image Retrieval, *Yi Xie, Huaidong Zhang, Xuemiao Xu, Jianqing Zhu, Shengfeng He*
349. CafeBoost: Causal Feature Boost to Eliminate Task-Induced Bias for Class Incremental Learning, *Benliu Qiu, Hongliang Li, Haitao Wen, Heqian Qiu, Lanxiao Wang, Fanman Meng, Qingbo Wu, Lili Pan*
350. Bilateral Memory Consolidation for Continual Learning, *Xing Nie, Shixiong Xu, Xiyan Liu, Gaofeng Meng, Chunlei Huo, Shiming Xiang*
351. NICO++: Towards Better Benchmarking for Domain Generalization, *Xingxuan Zhang, Yue He, Renzhe Xu, Han Yu, Zheyang Shen, Peng Cui*
352. DART: Diversify-Aggregate-Repeat Training Improves Generalization of Neural Networks, *Samyak Jain, Sravanti Addepalli, Pawan Kumar Sahu, Priyam Dey, R. Venkatesh Babu*
353. Differentiable Architecture Search With Random Features, *Xuanyang Zhang, Yonggang Li, Xiangyu Zhang, Yongtao Wang, Jian Sun*
354. Class Adaptive Network Calibration, *Bingyuan Liu, Jérôme Rony, Adrian Galdran, Jose Dolz, Ismail Ben Ayed*
355. Meta-Learning With a Geometry-Adaptive Preconditioner, *Suhyun Kang, Duhun Hwang, Moonjung Eo, Taesup Kim, Wonjong Rhee*
356. DepGraph: Towards Any Structural Pruning, *Gongfan Fang, Xinyin Ma, Mingli Song, Michael Bi Mi, Xinchao Wang*
357. Stitchable Neural Networks, *Zizheng Pan, Jianfei Cai, Bohan Zhuang*
358. Integral Neural Networks, *Kirill Solodskikh, Azim Kurbanov, Ruslan Aydarkhanov, Irina Zhelavskaya, Yury Parfenov, Dehua Song, Stamatios Lefkimmiatis*
359. Regularization of Polynomial Networks for Image Recognition, *Grigorios G. Chrysos, Bohan Wang, Jiankang Deng, Volkan Cevher*
360. ConvNeXt V2: Co-Designing and Scaling ConvNets With Masked Autoencoders, *Sanghyun Woo, Shoubhik Debnath, Ronghang Hu, Xinlei Chen, Zhuang Liu, In So Kweon, Saining Xie*
361. Shortcomings of Top-Down Randomization-Based Sanity Checks for Evaluations of Deep Neural Network Explanations, *Alexander Binder, Leander Weber, Sebastian Lapuschkin, Grégoire Montavon, Klaus-Robert Müller, Wojciech Samek*
362. Don't Lie to Me! Robust and Efficient Explainability With Verified Perturbation Analysis, *Thomas Fel, Melanie Ducoffe, David Vigouroux, Rémi Cadène, Mikaël Capelle, Claire Nicodème, Thomas Serre*
363. OT-Filter: An Optimal Transport Filter for Learning With Noisy Labels, *Chuanwen Feng, Yilong Ren, Xike Xie*
364. Robust Generalization Against Photon-Limited Corruptions via Worst-Case Sharpness Minimization, *Zhuo Huang, Miaoxi Zhu, Xiaobo Xia, Li Shen, Jun Yu, Chen Gong, Bo Han, Bo Du, Tongliang Liu*
365. Learning With Noisy Labels via Self-Supervised Adversarial Noisy Masking, *Yuanpeng Tu, Boshen Zhang, Yuxi Li, Liang Liu, Jian Li, Jiangning Zhang, Yabiao Wang, Chengjie Wang, Cai Rong Zhao*
366. Bit-Shrinking: Limiting Instantaneous Sharpness for Improving Post-Training Quantization, *Chen Lin, Bo Peng, Zheyang Li, Wenming Tan, Ye Ren, Jun Xiao, Shiliang Pu*
367. Enhancing Multiple Reliability Measures via Nuisance-Extended Information Bottleneck, *Jongheon Jeong, Sihyun Yu, Hankook Lee, Jinwoo Shin*
368. AdaptiveMix: Improving GAN Training via Feature Space Shrinkage, *Haozhe Liu, Wentian Zhang, Bing Li, Haoqian Wu, Nanjun He, Yawen Huang, Yuexiang Li, Bernard Ghanem, Yefeng Zheng*
369. Re-GAN: Data-Efficient GANs Training via Architectural Reconfiguration, *Divya Saxena, Jiannong Cao, Jiahao Xu, Tarun Kulshrestha*
370. Soft Augmentation for Image Classification, *Yang Liu, Shen Yan, Laura Leal-Taixé, James Hays, Deva Ramanan*
371. Boosting Verified Training for Robust Image Classifications via Abstraction, *Zhaodi Zhang, Zhiyi Xue, Yang Chen, Si Liu, Yueling Zhang, Jing Liu, Min Zhang*
372. A New Dataset Based on Images Taken by Blind People for Testing the Robustness of Image Classification Models Trained for ImageNet Categories, *Reza Akbarian Bafghi, Danna Gurari*
373. Exploiting Completeness and Uncertainty of Pseudo Labels for Weakly Supervised Video Anomaly Detection, *Chen Zhang, Guorong Li, Yuankai Qi, Shuhui Wang, Laiyun Qing, Qingming Huang, Ming-Hsuan Yang*
374. Prototypical Residual Networks for Anomaly Detection and Localization, *Hui Zhang, Zuxuan Wu, Zheng Wang, Zheneng Chen, Yu-Gang Jiang*
375. Class Balanced Adaptive Pseudo Labeling for Federated Semi-Supervised Learning, *Ming Li, Qingli Li, Yan Wang*
376. Fair Federated Medical Image Segmentation via Client Contribution Estimation, *Meirui Jiang, Holger R. Roth, Wenqi Li, Dong Yang, Can Zhao, Vishwesh Nath, Daguang Xu, Qi Dou, Ziyue Xu*
377. Rethinking Federated Learning With Domain Shift: A Prototype View, *Wenke Huang, Mang Ye, Zekun Shi, He Li, Bo Du*
378. FedDM: Iterative Distribution Matching for Communication-Efficient Federated Learning, *Yuanhao Xiong, Ruochen Wang, Minhao Cheng, Felix Yu, Cho-Jui Hsieh*
379. Alias-Free Convnets: Fractional Shift Invariance via Polynomial Activations, *Hagay Michaeli, Tomer Michaeli, Daniel Soudry*
380. STDLens: Model Hijacking-Resilient Federated Learning for Object Detection, *Ka-Ho Chow, Ling Liu, Wenqi Wei, Fatih Ilhan, Yanzhao Wu*
381. Detecting Backdoors in Pre-Trained Encoders, *Shiwei Feng, Guanhong Tao, Siyuan Cheng, Guangyu Shen, Xiangzhe Xu, Yingqi Liu, Kaiyuan Zhang, Shiqing Ma, Xiangyu Zhang*
382. Detecting Backdoors During the Inference Stage Based on Corruption Robustness Consistency, *Xiaogeng Liu, Minghui Li, Haoyu Wang, Shengshan Hu, Dengpan Ye, Hai Jin, Libing Wu, Chaowei Xiao*
383. Can't Steal? Cont-Steal! Contrastive Stealing Attacks Against Image Encoders, *Zeyang Sha, Xinlei He, Ning Yu, Michael Backes, Yang Zhang*

-
- A full-page sheet of white graph paper featuring a uniform grid of thin black horizontal and vertical lines. The grid covers the entire area of the page, providing a template for drawing or writing.

A full-page sheet of white graph paper featuring a uniform grid of thin gray lines. The grid consists of small squares covering the entire area, typical of standard graph paper used for mathematics or engineering.

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, creating a total area of 400 small squares. The margins are uniform on all sides.

Thursday, June 22

0800–1400 Registration (West Ballroom Foyer)**0800–0900 Breakfast** (West Ballrooms A–D)**0800–0900 Poster Setup** (West Exhibit Hall)**0900–1000 Plenary-Thu-AM** (East Exhibit Halls A–B)**Keynote:** Larry Zitnick (*Meta*)**Title:** Modeling Atoms to Address Our Climate Crisis

Abstract: Climate change is a societal and political problem whose impact could be mitigated by technology. Underlying many technical challenges is a surprisingly simple yet challenging problem; modeling the interaction of atoms. Approaching this problem from the perspective of a computer vision researcher has the potential to offer new insights into this growing and impactful field.

Chair: Ross Girshick (*Facebook*)**1000–1800 Exhibits** (West Exhibit Hall)

- See Exhibits map for list of exhibitors.

1000–1230 Demos (West Exhibit Hall Demo Area)

- Live 3D Reconstruction Using Stereo Event Cameras, *Suman Ghosh, Guillermo Gallego*
- Low-Cost Coded-Exposure-Pixel Cameras for Robust High-Speed Computational Imaging at Up to 18,000 Exposures-per-Second, *Roberto Rangel, Ayandev Barman, Rahul Gulve, Xiaonong Sun, Mian Wei, Don Nguyen, Motasem A. Sakr, David B. Lindell, Kiriakos N. Kutulakos, Roman Genov*
- Nothing Stands Still: A Spatiotemporal Benchmark on 3D Point Cloud Registration Under Large Geometric and Temporal Change, *Iro Armeni, Yan Hao, Shengyu Huang, Marc Pollefeys, Konrad Schindler, Tao Sun*
- Object-Aided SLAM in a World of Objects, *Matthieu Zins, Gilles Simon, Marie-Odile Berger, Romain Boisseau, Vincent Gaudillière*
- Open-Vocabulary Mobile Pick-and-Place, *Tsung-Yen Yang, Sergio Arnaud, Roozbeh Mottaghi, Oleksandr Maksymets, Mrinal Kalakrishnan, Eric Undersander, Dhruv Batra, Alexander William Clegg, Akshara Rai*
- PAniC-3D: 2D Anime Portraits to 3D Characters, *Shuhong Chen*
- Passive Micron-Scale Time-of-Flight With Sunlight Interferometry, *Alankar Kotwal, Anat Levin, Ioannis Gkioulekas*
- Photorealistic NeRF Models in Unreal Engine 5 As 3D MP4 Video, *Alex Grona*
- PyPose: A Library for Robot Learning With Physics-Based Optimization, *Chen Wang, Junyi Geng, Yaoyu Hu*
- Real-Time Neural Light Field on Mobile Devices, *Junli Cao, Pavlo Chemerys, Vladislav Shakhrai, Ju Hu, Sergey Tulyakov, Jian Ren*
- Real-Time Red Blood Cell Segmentation Using the Octopi Microscope Platform, *Kevin Marx*
- Real-Time Video-Based Stress Monitoring on Mobile Device, *Peyman Bateni, Leonid Sigal*

- RopeNet: Finetuning RepNet for Superhuman Scoring of Competitive Jump Rope Events, *Dylan Plummer*
- Scamp-7 Vision Sensor: Enabling High Speed In-Pixel Processing, *Laurie Bose, Piotr Dudek, Stephen Carey, Jianing Chen*
- SDFStudio: A Unified Framework for Surface Reconstruction, *Zehao Yu*
- MAGViT: Masked Generative Video Transformer, *Lijun Yu, Yong Cheng, Kihyuk Sohn, José Lezama, Han Zhang, Huiwen Chang, Alexander G. Hauptmann, Ming-Hsuan Yang, Yuan Hao, Irfan Essa, Lu Jiang*
- ARTificial Expressions: Interactive Human-Robot Painting, *Yejin Kim*
- Segment Anything, *Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alex Berg, Wan-Yen Lo, Piotr Dollar, Ross Girshick*

1030–1100 Morning Break (West Exhibit Hall)**1030–1230 Poster-Thu-AM** (West Exhibit Hall)

✧ - Highlight paper (check it out)

☞ - Award candidate paper (see award sessions)

1. Fresnel Microfacet BRDF: Unification of Polari-Radiometric Surface-Body Reflection, *Tomoki Ichikawa, Yoshiki Fukao, Shohei Nobuhara, Ko Nishino*
2. JacobiNeRF: NeRF Shaping With Mutual Information Gradients, *Xiaomeng Xu, Yanchao Yang, Kaichun Mo, Boxiao Pan, Li Yi, Leonidas Guibas*
3. ContraNeRF: Generalizable Neural Radiance Fields for Synthetic-to-Real Novel View Synthesis via Contrastive Learning, *Hao Yang, Lanqing Hong, Aoxue Li, Tianyang Hu, Zhenguo Li, Gim Hee Lee, Liwei Wang*
4. SCADE: NeRFs from Space Carving With Ambiguity-Aware Depth Estimates, *Mikaela Angelina Uy, Ricardo Martin-Brualla, Leonidas Guibas, Ke Li*
5. Removing Objects From Neural Radiance Fields, *Silvan Weder, Guillermo Garcia-Hernando, Áron Monszpart, Marc Pollefeys, Gabriel J. Brostow, Michael Firman, Sara Vicente*
6. Progressively Optimized Local Radiance Fields for Robust View Synthesis, *Andreas Meuleman, Yu-Lun Liu, Chen Gao, Jia-Bin Huang, Changil Kim, Min H. Kim, Johannes Kopf*
7. NeRFVS: Neural Radiance Fields for Free View Synthesis via Geometry Scaffolds, *Chen Yang, Peihao Li, Zanwei Zhou, Shanxin Yuan, Bingbing Liu, Xiaokang Yang, Weichao Qiu, Wei Shen*
8. ABLE-NeRF: Attention-Based Rendering With Learnable Embeddings for Neural Radiance Field, *Zhe Jun Tang, Tat-Jen Cham, Haiyu Zhao*
9. MobileNeRF: Exploiting the Polygon Rasterization Pipeline for Efficient Neural Field Rendering on Mobile Architectures, *Zhiqin Chen, Thomas Funkhouser, Peter Hedman, Andrea Tagliasacchi*
10. pCON: Polarimetric Coordinate Networks for Neural Scene Representations, *Henry Peters, Yunhao Ba, Achuta Kadambi*
11. Balanced Spherical Grid for Egocentric View Synthesis, *Changwoon Choi, Sang Min Kim, Young Min Kim*
12. Complementary Intrinsic From Neural Radiance Fields and CNNs for Outdoor Scene Relighting, *Siqi Yang, Xuanning Cui, Yongjie Zhu, Jiajun Tang, Si Li, Zhaoqi Yu, Boxin Shi*

13. HyperReel: High-Fidelity 6-DoF Video With Ray-Conditioned Sampling, *Benjamin Attal, Jia-Bin Huang, Christian Richardt, Michael Zollhöfer, Johannes Kopf, Matthew O'Toole, Changil Kim*
14. UV Volumes for Real-Time Rendering of Editable Free-View Human Performance, *Yue Chen, Xuan Wang, Xingyu Chen, Qi Zhang, Xiaoyu Li, Yu Guo, Jue Wang, Fei Wang*
15. Tensor4D: Efficient Neural 4D Decomposition for High-Fidelity Dynamic Reconstruction and Rendering, *Ruizhi Shao, Zerong Zheng, Hanzhang Tu, Boning Liu, Hongwen Zhang, Yebin Liu*
16. PixHt-Lab: Pixel Height Based Light Effect Generation for Image Compositing, *Yichen Sheng, Jianming Zhang, Julien Philip, Yannick Hold-Geoffroy, Xin Sun, He Zhang, Lu Ling, Bedrich Benes*
17. Computational Flash Photography Through Intrinsic, *Sepideh Sarajian Maralan, Chris Careaga, Yagiz Aksoy*
18. RelightableHands: Efficient Neural Relighting of Articulated Hand Models, *Shun Iwase, Shunsuke Saito, Tomas Simon, Stephen Lombardi, Timur Bagautdinov, Rohan Joshi, Fabian Prada, Takaaki Shiratori, Yaser Sheikh, Jason Saragih*
19. TMO: Textured Mesh Acquisition of Objects With a Mobile Device by Using Differentiable Rendering, *Jaehoon Choi, Dongki Jung, Taejae Lee, Sangwook Kim, Youngdong Jung, Dinesh Manocha, Donghwan Lee*
20. VolRecon: Volume Rendering of Signed Ray Distance Functions for Generalizable Multi-View Reconstruction, *Yufan Ren, Fangjinhua Wang, Tong Zhang, Marc Pollefeys, Sabine Süssstrunk*
21. Multi-View Reconstruction Using Signed Ray Distance Functions (SRDF), *Pierre Zins, Yuanlu Xu, Edmond Boyer, Stefanie Wuhler, Tony Tung*
22. Structural Multiplane Image: Bridging Neural View Synthesis and 3D Reconstruction, *Mingfang Zhang, Jinglu Wang, Xiao Li, Yifei Huang, Yoichi Sato, Yan Lu*
23. Octree Guided Unoriented Surface Reconstruction, *Chamin Hewa Koneputugodage, Yizhak Ben-Shabat, Stephen Gould*
24. Neural Vector Fields: Implicit Representation by Explicit Learning, *Xianghui Yang, Guosheng Lin, Zhenghao Chen, Luping Zhou*
25. DA Wand: Distortion-Aware Selection Using Neural Mesh Parameterization, *Richard Liu, Noam Aigerman, Vladimir G. Kim, Rana Hanocka*
26. Diffusion-Based Generation, Optimization, and Planning in 3D Scenes, *Siyuan Huang, Zan Wang, Puhao Li, Baoxiong Jia, Tengyu Liu, Yixin Zhu, Wei Liang, Song-Chun Zhu*
27. Patch-Based 3D Natural Scene Generation From a Single Example, *Weiyu Li, Xuelin Chen, Jue Wang, Baoquan Chen*
28. Consistent View Synthesis With Pose-Guided Diffusion Models, *Hung-Yu Tseng, Qinbo Li, Changil Kim, Suhub Alsisan, Jia-Bin Huang, Johannes Kopf*
29. Generalized Deep 3D Shape Prior via Part-Discretized Diffusion Process, *Yuhan Li, Yishun Dou, Xuanhong Chen, Bingbing Ni, Yilin Sun, Yutian Liu, Fuzhen Wang*
30. High Fidelity 3D Hand Shape Reconstruction via Scalable Graph Frequency Decomposition, *Tianyu Luan, Yuanhao Zhai, Jingjing Meng, Zhong Li, Zhang Chen, Yi Xu, Junsong Yuan*
31. TAPS3D: Text-Guided 3D Textured Shape Generation From Pseudo Supervision, *Jiacheng Wei, Hao Wang, Jiashi Feng, Guosheng Lin, Kim-Hui Yap*
32. SECAD-Net: Self-Supervised CAD Reconstruction by Learning Sketch-Extrude Operations, *Pu Li, Jianwei Guo, Xiaopeng Zhang, Dong-Ming Yan*
33. Interactive Cartoonization With Controllable Perceptual Factors, *Namhyuk Ahn, Patrick Kwon, Jihye Back, Kibeom Hong, Seungkwon Kim*
34. High-Res Facial Appearance Capture From Polarized Smartphone Images, *Dejan Azinović, Olivier Maury, Christophe Hery, Matthias Nießner, Justus Thies*
35. GlassesGAN: Eyewear Personalization Using Synthetic Appearance Discovery and Targeted Subspace Modeling, *Richard Plesh, Peter Peer, Vitomir Struc*
36. Continuous Landmark Detection With 3D Queries, *Prashanth Chandran, Gaspard Zoss, Paulo Gotardo, Derek Bradley*
37. NeuFace: Realistic 3D Neural Face Rendering From Multi-View Images, *Mingwu Zheng, Haiyu Zhang, Hongyu Yang, Di Huang*
38. AVFace: Towards Detailed Audio-Visual 4D Face Reconstruction, *Aggelina Chatziagapi, Dimitris Samaras*
39. Learning Personalized High Quality Volumetric Head Avatars From Monocular RGB Videos, *Ziqian Bai, Feitong Tan, Zeng Huang, Kripasindhu Sarkar, Danhang Tang, Di Qiu, Abhimitra Meka, Ruofei Du, Mingsong Dou, Sergio Orts-Escolano, Rohit Pandey, Ping Tan, Thabo Beeler, Sean Fanello, Yinda Zhang*
40. OTAvatar: One-Shot Talking Face Avatar With Controllable Tri-Plane Rendering, *Zhiyuan Ma, Xiangyu Zhu, Guo-Jun Qi, Zhen Lei, Lei Zhang*
41. X-Avatar: Expressive Human Avatars, *Kaiyue Shen, Chen Guo, Manuel Kaufmann, Juan Jose Zarate, Julien Valentin, Jie Song, Otmar Hilliges*
42. InstantAvatar: Learning Avatars From Monocular Video in 60 Seconds, *Tianjian Jiang, Xu Chen, Jie Song, Otmar Hilliges*
43. JAWS: Just a Wild Shot for Cinematic Transfer in Neural Radiance Fields, *Xi Wang, Robin Courant, Jinglei Shi, Eric Marchand, Marc Christie*
44. MonoHuman: Animatable Human Neural Field From Monocular Video, *Zhengming Yu, Wei Cheng, Xian Liu, Wayne Wu, Kwan-Yee Lin*
45. Structured 3D Features for Reconstructing Controllable Avatars, *Enric Corona, Mihai Zanfir, Thiemo Alldieck, Eduard Gabriel Bazavan, Andrei Zanfir, Cristian Sminchisescu*
46. HOOD: Hierarchical Graphs for Generalized Modelling of Clothing Dynamics, *Artur Grigorev, Michael J. Black, Otmar Hilliges*
47. Physically Realizable Natural-Looking Clothing Textures Evade Person Detectors via 3D Modeling, *Zhanhao Hu, Wenda Chu, Xiaopei Zhu, Hui Zhang, Bo Zhang, Xiaolin Hu*
48. Learning Semantic-Aware Disentangled Representation for Flexible 3D Human Body Editing, *Xiaokun Sun, Qiao Feng, Xiongzheng Li, Jinsong Zhang, Yu-Kun Lai, Jingyu Yang, Kun Li*
49. Reconstructing Animatable Categories From Videos, *Gengshan Yang, Chaoyang Wang, N. Dinesh Reddy, Deva Ramanan*
50. Deformable Mesh Transformer for 3D Human Mesh Recovery, *Yusuke Yoshiyasu*
51. Hi4D: 4D Instance Segmentation of Close Human Interaction, *Yifei Yin, Chen Guo, Manuel Kaufmann, Juan Jose Zarate, Jie Song, Otmar Hilliges*
52. Bringing Inputs to Shared Domains for 3D Interacting Hands Recovery in the Wild, *Gyeongsik Moon*

53. Learning Human Mesh Recovery in 3D Scenes, *Zehong Shen, Zhi Cen, Sida Peng, Qing Shuai, Hujun Bao, Xiaowei Zhou*
54. H2ONet: Hand-Occlusion-and-Orientation-Aware Network for Real-Time 3D Hand Mesh Reconstruction, *Hao Xu, Tianyu Wang, Xiao Tang, Chi-Wing Fu*
55. What You Can Reconstruct From a Shadow, *Ruoshi Liu, Sachit Menon, Chengzhi Mao, Dennis Park, Simon Stent, Carl Vondrick*
56. Autonomous Manipulation Learning for Similar Deformable Objects via Only One Demonstration, *Yu Ren, Ronghan Chen, Yang Cong*
57. In-Hand 3D Object Scanning From an RGB Sequence, *Shreyas Hampali, Tomas Hodan, Luan Tran, Lingni Ma, Cem Keskin, Vincent Lepetit*
58. Putting People in Their Place: Affordance-Aware Human Insertion Into Scenes, *Sumith Kulal, Tim Brooks, Alex Aiken, Jiajun Wu, Jimei Yang, Jingwan Lu, Alexei A. Efros, Krishna Kumar Singh*
59. Detecting Human-Object Contact in Images, *Yixin Chen, Sai Kumar Dwivedi, Michael J. Black, Dimitrios Tzionas*
60. What Happened 3 Seconds Ago? Inferring the Past With Thermal Imaging, *Zitian Tang, Wenjie Ye, Wei-Chiu Ma, Hang Zhao*
61. Trajectory-Aware Body Interaction Transformer for Multi-Person Pose Forecasting, *Xiaogang Peng, Siyuan Mao, Zizhao Wu*
62. Mutual Information-Based Temporal Difference Learning for Human Pose Estimation in Video, *Runyang Feng, Yixing Gao, Xueqing Ma, Tze Ho Elden Tse, Hyung Jin Chang*
63. Ego-Body Pose Estimation via Ego-Head Pose Estimation, *Jiaman Li, Karen Liu, Jiajun Wu*
64. ViPLO: Vision Transformer Based Pose-Conditioned Self-Loop Graph for Human-Object Interaction Detection, *Jeeseung Park, Jin-Woo Park, Jong-Seok Lee*
65. HS-Pose: Hybrid Scope Feature Extraction for Category-Level Object Pose Estimation, *Linfang Zheng, Chen Wang, Yinghan Sun, Esha Dasgupta, Hua Chen, Aleš Leonardis, Wei Zhang, Hyung Jin Chang*
66. ScarceNet: Animal Pose Estimation With Scarce Annotations, *Chen Li, Gim Hee Lee*
67. Cross-Domain 3D Hand Pose Estimation With Dual Modalities, *Qiuxia Lin, Linlin Yang, Angela Yao*
68. Linking Garment With Person via Semantically Associated Landmarks for Virtual Try-On, *Keyu Yan, Tingwei Gao, Hui Zhang, Chengjun Xie*
69. Level-S²fM: Structure From Motion on Neural Level Set of Implicit Surfaces, *Yuxi Xiao, Nan Xue, Tianfu Wu, Gui-Song Xia*
70. Revisiting Rotation Averaging: Uncertainties and Robust Losses, *Ganlin Zhang, Viktor Larsson, Daniel Barath*
71. SliceMatch: Geometry-Guided Aggregation for Cross-View Pose Estimation, *Ted Lentsch, Zimin Xia, Holger Caesar, Julian F. P. Kooij*
72. Learning the Distribution of Errors in Stereo Matching for Joint Disparity and Uncertainty Estimation, *Liyan Chen, Weihan Wang, Philippos Mordohai*
73. Long-Term Visual Localization With Mobile Sensors, *Shen Yan, Yu Liu, Long Wang, Zehong Shen, Zhen Peng, Haomin Liu, Maojun Zhang, Guofeng Zhang, Xiaowei Zhou*
74. Learning to Predict Scene-Level Implicit 3D From Posed RGBD Data, *Nilesh Kulkarni, Linyi Jin, Justin Johnson, David F. Fouhey*
75. Paired-Point Lifting for Enhanced Privacy-Preserving Visual Localization, *Chunghwan Lee, Jaihoon Kim, Chanhyuk Yun, Je Hyeong Hong*
76. The ObjectFolder Benchmark: Multisensory Learning With Neural and Real Objects, *Ruohan Gao, Yiming Dou, Hao Li, Tanmay Agarwal, Jeannette Bohg, Yunzhu Li, Li Fei-Fei, Jiajun Wu*
77. Learning Accurate 3D Shape Based on Stereo Polarimetric Imaging, *Tianyu Huang, Haoang Li, Kejing He, Congying Sui, Bin Li, Yun-Hui Liu*
78. RUST: Latent Neural Scene Representations From Unposed Imagery, *Mehdi S. M. Sajjadi, Aravindh Mahendran, Thomas Kipf, Etienne Pot, Daniel Duckworth, Mario Lučić, Klaus Greff*
79. Perspective Fields for Single Image Camera Calibration, *Linyi Jin, Jianming Zhang, Yannick Hold-Geoffroy, Oliver Wang, Kevin Blackburn-Matzen, Matthew Sticha, David F. Fouhey*
80. VisFusion: Visibility-Aware Online 3D Scene Reconstruction From Videos, *Huiyu Gao, Wei Mao, Miaomiao Liu*
81. DeepLSD: Line Segment Detection and Refinement With Deep Image Gradients, *Rémi Pautrat, Daniel Barath, Viktor Larsson, Martin R. Oswald, Marc Pollefeys*
82. Disentangling Orthogonal Planes for Indoor Panoramic Room Layout Estimation With Cross-Scale Distortion Awareness, *Zhijie Shen, Zishuo Zheng, Chunyu Lin, Lang Nie, Kang Liao, Shuai Zheng, Yao Zhao*
83. Single Image Depth Prediction Made Better: A Multivariate Gaussian Take, *Ce Liu, Suryansh Kumar, Shuhang Gu, Radu Timofte, Luc Van Gool*
84. Wide-Angle Rectification via Content-Aware Conformal Mapping, *Qi Zhang, Hongdong Li, Qing Wang*
85. All-in-Focus Imaging From Event Focal Stack, *Hanyue Lou, Minggui Teng, Yixin Yang, Boxin Shi*
86. Multi-View Stereo Representation Revist: Region-Aware MVSNet, *Yisu Zhang, Jianke Zhu, Lixiang Lin*
87. Semantic Ray: Learning a Generalizable Semantic Field With Cross-Reprojection Attention, *Fangfu Liu, Chubin Zhang, Yu Zheng, Yueqi Duan*
88. OmniCity: Omnipotent City Understanding With Multi-Level and Multi-View Images, *Weijia Li, Yawen Lai, Linning Xu, Yuanbo Xiangli, Jinhua Yu, Conghui He, Gui-Song Xia, Dahua Lin*
89. ESLAM: Efficient Dense SLAM System Based on Hybrid Representation of Signed Distance Fields, *Mohammad Mahdi Johari, Camilla Carta, François Fleuret*
90. Non-Line-of-Sight Imaging With Signal Superresolution Network, *Jianyu Wang, Xintong Liu, Leping Xiao, Zuoqiang Shi, Lingyun Qiu, Xing Fu*
91. Look, Radiate, and Learn: Self-Supervised Localisation via Radio-Visual Correspondence, *Mohammed Alloulah, Maximilian Arnold*
92. Learning Transformations to Reduce the Geometric Shift in Object Detection, *Vidit Vidit, Martin Engilberge, Mathieu Salzmann*
93. Anchor3DLane: Learning to Regress 3D Anchors for Monocular 3D Lane Detection, *Shaofei Huang, Zhenwei Shen, Zehao Huang, Zi-han Ding, Jiao Dai, Jizhong Han, Naiyan Wang, Si Liu*
94. BEV-SAN: Accurate BEV 3D Object Detection via Slice Attention Networks, *Xiaowei Chi, Jiaming Liu, Ming Lu, Rongyu Zhang, Zhaoqing Wang, Yandong Guo, Shanghang Zhang*
95. Semi-Supervised Stereo-Based 3D Object Detection via Cross-View Consensus, *Wenhao Wu, Hau San Wong, Si Wu*

96. Weakly Supervised Monocular 3D Object Detection Using Multi-View Projection and Direction Consistency, *Runzhou Tao, Wencheng Han, Zhongying Qiu, Cheng-Zhong Xu, Jianbing Shen*
97. MonoATT: Online Monocular 3D Object Detection With Adaptive Token Transformer, *Yunsong Zhou, Hongzi Zhu, Quan Liu, Shan Chang, Minyi Guo*
98. Azimuth Super-Resolution for FMCW Radar in Autonomous Driving, *Yu-Jhe Li, Shawn Hunt, Jinhyung Park, Matthew O'Toole, Kris Kitani*
99. Pix2map: Cross-Modal Retrieval for Inferring Street Maps From Images, *Xindi Wu, KwunFung Lau, Francesco Ferroni, Aljoša Ošep, Deva Ramanan*
100. LoGoNet: Towards Accurate 3D Object Detection With Local-to-Global Cross-Modal Fusion, *Xin Li, Tao Ma, Yuenan Hou, Botian Shi, Yuchen Yang, Youquan Liu, Xingjiao Wu, Qin Chen, Yikang Li, Yu Qiao, Liang He*
101. Neural Map Prior for Autonomous Driving, *Xuan Xiong, Yicheng Liu, Tianyuan Yuan, Yue Wang, Yilun Wang, Hang Zhao*
102. Spherical Transformer for LiDAR-Based 3D Recognition, *Xin Lai, Yukang Chen, Fanbin Lu, Jianhui Liu, Jiaya Jia*
103. Density-Insensitive Unsupervised Domain Adaption on 3D Object Detection, *Qianjiang Hu, Daizong Liu, Wei Hu*
104. PillarNeXt: Rethinking Network Designs for 3D Object Detection in LiDAR Point Clouds, *Jinyu Li, Chenxu Luo, Xiaodong Yang*
105. PeakConv: Learning Peak Receptive Field for Radar Semantic Segmentation, *Liwen Zhang, Xinyan Zhang, Youcheng Zhang, Yufei Guo, Yuanpei Chen, Xuhui Huang, Zhe Ma*
106. Single Domain Generalization for LiDAR Semantic Segmentation, *Hyeonseong Kim, Yoonsu Kang, Changgyoon Oh, Kuk-Jin Yoon*
107. Weakly Supervised Class-Agnostic Motion Prediction for Autonomous Driving, *Ruibao Li, Hanyu Shi, Ziang Fu, Zhe Wang, Guosheng Lin*
108. MethaneMapper: Spectral Absorption Aware Hyperspectral  Transformer for Methane Detection, *Satish Kumar, Ivan Arevalo, ASM Iftekhar, B S Manjunath*
109. GrowSP: Unsupervised Semantic Segmentation of 3D Point Clouds, *Zihui Zhang, Bo Yang, Bing Wang, Bo Li*
110. SCoDA: Domain Adaptive Shape Completion for Real Scans, *Yushuang Wu, Zizheng Yan, Ce Chen, Lai Wei, Xiao Li, Guanbin Li, Yihao Li, Shuguang Cui, Xiaoguang Han*
111. SCPNet: Semantic Scene Completion on Point Cloud, *Zhaoyang  Xia, Youquan Liu, Xin Li, Xinge Zhu, Yuexin Ma, Yikang Li, Yuenan Hou, Yu Qiao*
112. ViewNet: A Novel Projection-Based Backbone With View Pooling for Few-Shot Point Cloud Classification, *Jiajing Chen, Minmin Yang, Senem Velipasalar*
113. Complete-to-Partial 4D Distillation for Self-Supervised Point Cloud Sequence Representation Learning, *Zhuoyang Zhang, Yuhao Dong, Yunze Liu, Li Yi*
114. Learnable Skeleton-Aware 3D Point Cloud Sampling, *Cheng Wen, Baosheng Yu, Dacheng Tao*
115. Meta Architecture for Point Cloud Analysis, *Haojia Lin, Xiawu Zheng, Lijiang Li, Fei Chao, Shanshan Wang, Yan Wang, Yonghong Tian, Rongrong Ji*
116. PointListNet: Deep Learning on 3D Point Lists, *Hehe Fan, Linchao Zhu, Yi Yang, Mohan Kankanhalli*
117. PEAL: Prior-Embedded Explicit Attention Learning for Low-Overlap Point Cloud Registration, *Junle Yu, Luwei Ren, Yu Zhang, Wenhui Zhou, Lili Lin, Guojun Dai*
118. Unsupervised Inference of Signed Distance Functions From Single Sparse Point Clouds Without Learning Priors, *Chao Chen, Yu-Shen Liu, Zhizhong Han*
119. Towards Better Gradient Consistency for Neural Signed Distance Functions via Level Set Alignment, *Baorui Ma, Junsheng Zhou, Yu-Shen Liu, Zhizhong Han*
120. Self-Supervised Learning for Multimodal Non-Rigid 3D Shape  Matching, *Dongliang Cao, Florian Bernard*
121. 3D Registration With Maximal Cliques, *Xiyu Zhang, Jiaqi Yang,  Shikun Zhang, Yanning Zhang*
122. PanoSwin: A Pano-Style Swin Transformer for Panorama Understanding, *Zhixin Ling, Zhen Xing, Xiangdong Zhou, Manliang Cao, Guichun Zhou*
123. DKM: Dense Kernelized Feature Matching for Geometry  Estimation, *Johan Edstedt, Ioannis Athanasiadis, Mårten Wadenbäck, Michael Felsberg*
124. PATS: Patch Area Transportation With Subdivision for Local Feature Matching, *Junjie Ni, Yijin Li, Zhaoyang Huang, Hongsheng Li, Hujun Bao, Zhaopeng Cui, Guofeng Zhang*
125. Correspondence Transformers With Asymmetric Feature Learning and Matching Flow Super-Resolution, *Yixuan Sun, Dongyang Zhao, Zhangyue Yin, Yiwen Huang, Tao Gui, Wenqiang Zhang, Weifeng Ge*
126. Learning Adaptive Dense Event Stereo From the Image Domain, *Hoonhee Cho, Jegyeong Cho, Kuk-Jin Yoon*
127. On the Convergence of IRLS and Its Variants in Outlier-Robust  Estimation, *Liangzu Peng, Christian Kümmeler, René Vidal*
128. You Only Segment Once: Towards Real-Time Panoptic Segmentation, *Jie Hu, Linyan Huang, Tianhe Ren, Shengchuan Zhang, Rongrong Ji, Liujuan Cao*
129. BEVFormer v2: Adapting Modern Image Backbones to Bird's-  Eye-View Recognition via Perspective Supervision, *Chenyu Yang, Yuntao Chen, Hao Tian, Chenxin Tao, Xizhou Zhu, Zhaoxiang Zhang, Gao Huang, Hongyang Li, Yu Qiao, Lewei Lu, Jie Zhou, Jifeng Dai*
130. UniHCP: A Unified Model for Human-Centric Perceptions, *Yuanzheng Ci, Yizhou Wang, Meilin Chen, Shixiang Tang, Lei Bai, Feng Zhu, Rui Zhao, Fengwei Yu, Donglian Qi, Wanli Ouyang*
131. Planning-Oriented Autonomous Driving, *Yihan Hu, Jiazhi Yang,  Li Chen, Keyu Li, Chonghao Sima, Xizhou Zhu, Siqi Chai, Senyao Du, Tianwei Lin, Wenhai Wang, Lewei Lu, Xiaosong Jia, Qiang Liu, Jifeng Dai, Yu Qiao, Hongyang Li*
132. Query-Centric Trajectory Prediction, *Zikang Zhou, Jianping Wang, Yung-Hui Li, Yu-Kai Huang*
133. Unsupervised Sampling Promoting for Stochastic Human Trajectory Prediction, *Guangyi Chen, Zhenhao Chen, Shunxing Fan, Kun Zhang*
134. AdamsFormer for Spatial Action Localization in the Future, *Hyung-gun Chi, Kwonjoon Lee, Nakul Agarwal, Yi Xu, Karthik Ramani, Chihoh Choi*
135. PIRLNav: Pretraining With Imitation and RL Finetuning for ObjectNav, *Ram Ramrakhya, Dhruv Batra, Erik Wijmans, Abhishek Das*
136. NeRF in the Palm of Your Hand: Corrective Augmentation for Robotics via Novel-View Synthesis, *Allan Zhou, Moo Jin Kim, Lirui Wang, Pete Florence, Chelsea Finn*

137. Camouflaged Instance Segmentation via Explicit De-
✧ Camouflaging, *Naisong Luo, Yuwen Pan, Rui Sun, Tianzhu Zhang, Zhiwei Xiong, Feng Wu*
138. Standing Between Past and Future: Spatio-Temporal Modeling for Multi-Camera 3D Multi-Object Tracking, *Ziqi Pang, Jie Li, Pavel Tokmakov, Dian Chen, Sergey Zagoruyko, Yu-Xiong Wang*
139. MotionTrack: Learning Robust Short-Term and Long-Term Motions for Multi-Object Tracking, *Zheng Qin, Sanping Zhou, Le Wang, Jinghai Duan, Gang Hua, Wei Tang*
140. Multi-Modal Gait Recognition via Effective Spatial-Temporal Feature Fusion, *Yufeng Cui, Yimei Kang*
141. Rethinking the Learning Paradigm for Dynamic Facial Expression Recognition, *Hanyang Wang, Bo Li, Shuang Wu, Siyuan Shen, Feng Liu, Shouhong Ding, Aimin Zhou*
142. One-Shot High-Fidelity Talking-Head Synthesis With Deformable Neural Radiance Field, *Weichuang Li, Longhao Zhang, Dong Wang, Bin Zhao, Zhigang Wang, Mulin Chen, Bang Zhang, Zhongjian Wang, Liefeng Bo, Xuelong Li*
143. Progressive Disentangled Representation Learning for Fine-Grained Controllable Talking Head Synthesis, *Duomin Wang, Yu Deng, Zixin Yin, Heung-Yeung Shum, Baoyuan Wang*
144. Event-Guided Person Re-Identification via Sparse-Dense Complementary Learning, *Chengzhi Cao, Xueyang Fu, Hongjian Liu, Yukun Huang, Kunyu Wang, Jiebo Luo, Zheng-Jun Zha*
145. Executing Your Commands via Motion Diffusion in Latent Space, *Xin Chen, Biao Jiang, Wen Liu, Zilong Huang, Bin Fu, Tao Chen, Gang Yu*
146. MoLo: Motion-Augmented Long-Short Contrastive Learning for Few-Shot Action Recognition, *Xiang Wang, Shiwei Zhang, Zhiwu Qing, Changxin Gao, Yingya Zhang, Deli Zhao, Nong Sang*
147. "Seeing" Electric Network Frequency From Events, *Lexuan Xu, Guang Hua, Haijian Zhang, Lei Yu, Ning Qiao*
148. Event-Based Video Frame Interpolation With Cross-Modal
✧ Asymmetric Bidirectional Motion Fields, *Taewoo Kim, Yujeong Chae, Hyun-Kurl Jang, Kuk-Jin Yoon*
149. Event-Based Frame Interpolation With Ad-Hoc Deblurring, *Lei Sun, Christos Sakaridis, Jingyun Liang, Peng Sun, Jiezhong Cao, Kai Zhang, Qi Jiang, Kaiwei Wang, Luc Van Gool*
150. Video Dehazing via a Multi-Range Temporal Alignment Network With Physical Prior, *Jiaqi Xu, Xiaowei Hu, Lei Zhu, Qi Dou, Jifeng Dai, Yu Qiao, Pheng-Ann Heng*
151. TransFlow: Transformer As Flow Learner, *Yawen Lu, Qifan Wang, Siqi Ma, Tong Geng, Yingjie Victor Chen, Huaijin Chen, Dongfang Liu*
152. MP-Former: Mask-Piloted Transformer for Image Segmentation, *Hao Zhang, Feng Li, Huaizhe Xu, Shijia Huang, Shilong Liu, Lionel M. Ni, Lei Zhang*
153. GradICON: Approximate Diffeomorphisms via Gradient Inverse Consistency, *Lin Tian, Hastings Greer, François-Xavier Vialard, Roland Kwitt, Raúl San José Estépar, Richard Jarrett Rushmore, Nikolaos Makris, Sylvain Bouix, Marc Niethammer*
154. Neural Texture Synthesis With Guided Correspondence, *Yang Zhou, Kaijian Chen, Rongjun Xiao, Hui Huang*
155. Self-Supervised Non-Uniform Kernel Estimation With Flow-Based Motion Prior for Blind Image Deblurring, *Zhenxuan Fang, Fangfang Wu, Weisheng Dong, Xin Li, Jinjian Wu, Guangming Shi*
156. Decoupling-and-Aggregating for Image Exposure Correction, *Yang Wang, Long Peng, Liang Li, Yang Cao, Zheng-Jun Zha*
157. You Do Not Need Additional Priors or Regularizers in Retinex-Based Low-Light Image Enhancement, *Huiyuan Fu, Wenkai Zheng, Xiangyu Meng, Xin Wang, Chuanming Wang, Huadong Ma*
158. DNF: Decouple and Feedback Network for Seeing in the Dark, ✧ *Xin Jin, Ling-Hao Han, Zhen Li, Chun-Le Guo, Zhi Chai, Chongyi Li*
159. Contrastive Semi-Supervised Learning for Underwater Image Restoration via Reliable Bank, *Shirui Huang, Keyan Wang, Huan Liu, Jun Chen, Yunsong Li*
160. LG-BPN: Local and Global Blind-Patch Network for Self-Supervised Real-World Denoising, *Zichun Wang, Ying Fu, Ji Liu, Yulun Zhang*
161. Spectral Bayesian Uncertainty for Image Super-Resolution, *Tao Liu, Jun Cheng, Shan Tan*
162. Deep Random Projector: Accelerated Deep Image Prior, *Taihui Li, Hengkang Wang, Zhong Zhuang, Ju Sun*
163. Context-Aware Pretraining for Efficient Blind Image Decomposition, *Chao Wang, Zhedong Zheng, Ruijie Quan, Yifan Sun, Yi Yang*
164. Metadata-Based RAW Reconstruction via Implicit Neural Functions, *Leyi Li, Huijie Qiao, Qi Ye, Qinmin Yang*
165. Raw Image Reconstruction With Learned Compact Metadata, *Yufei Wang, Yi Yu, Wenhan Yang, Lanqing Guo, Lap-Pui Chau, Alex C. Kot, Bihan Wen*
166. AccelIR: Task-Aware Image Compression for Accelerating Neural Restoration, *Juncheol Ye, Hyunho Yeo, Jinwoo Park, Dongsu Han*
167. AutoFocusFormer: Image Segmentation off the Grid, *Chen Ziwen, Kaushik Patnaik, Shuangfei Zhai, Alvin Wan, Zhile Ren, Alexander G. Schwing, Alex Colburn, Li Fuxin*
168. Guided Depth Super-Resolution by Deep Anisotropic Diffusion, *Nando Metzger, Rodrigo Caye Daudt, Konrad Schindler*
169. Super-Resolution Neural Operator, *Min Wei, Xuesong Zhang*
170. Cascaded Local Implicit Transformer for Arbitrary-Scale Super-Resolution, *Hao-Wei Chen, Yu-Syuan Xu, Min-Fong Hong, Yi-Min Tsai, Hsien-Kai Kuo, Chun-Yi Lee*
171. GamutMLP: A Lightweight MLP for Color Loss Recovery, *Hoang M. Le, Brian Price, Scott Cohen, Michael S. Brown*
172. Efficient and Explicit Modelling of Image Hierarchies for Image Restoration, *Yawei Li, Yuchen Fan, Xiaoyu Xiang, Denis Demandolx, Rakesh Ranjan, Radu Timofte, Luc Van Gool*
173. LEMaRT: Label-Efficient Masked Region Transform for Image Harmonization, *Sheng Liu, Cong Phuoc Huynh, Cong Chen, Maxim Arap, Raffay Hamid*
174. CAP-VSTNet: Content Affinity Preserved Versatile Style Transfer, *Linfeng Wen, Chengying Gao, Changqing Zou*
175. ObjectStitch: Object Compositing With Diffusion Model, *Yizhi Song, Zhifei Zhang, Zhe Lin, Scott Cohen, Brian Price, Jianming Zhang, Soo Ye Kim, Daniel Aliaga*
176. DeepVecFont-v2: Exploiting Transformers to Synthesize Vector Fonts With Higher Quality, *Yuqing Wang, Yizhi Wang, Longhui Yu, Yuesheng Zhu, Zhouhui Lian*
177. Master: Meta Style Transformer for Controllable Zero-Shot and Few-Shot Artistic Style Transfer, *Hao Tang, Songhua Liu, Tianwei Lin, Shaoli Huang, Fu Li, Dongliang He, Xinchao Wang*
178. CLIP-Sculptor: Zero-Shot Generation of High-Fidelity and Diverse Shapes From Natural Language, *Aditya Sanghi, Rao Fu, Vivian Liu, Karl D.D. Willis, Hooman Shayan, Amir H. Khasahmadi, Srinath Sridhar, Daniel Ritchie*

179. LayoutDM: Transformer-Based Diffusion Model for Layout Generation, *Shang Chai, Liansheng Zhuang, Fengying Yan*
180. Imagen Editor and EditBench: Advancing and Evaluating Text-
✧ Guided Image Inpainting, *Su Wang, Chitwan Saharia, Ceslee Montgomery, Jordi Pont-Tuset, Shai Noy, Stefano Pellegrini, Yasumasa Onoe, Sarah Laszlo, David J. Fleet, Radu Soricut, Jason Baldridge, Mohammad Norouzi, Peter Anderson, William Chan*
181. SpaText: Spatio-Textual Representation for Controllable Image Generation, *Omri Avrahami, Thomas Hayes, Oran Gafni, Sonal Gupta, Yaniv Taigman, Devi Parikh, Dani Lischinski, Ohad Fried, Xi Yin*
182. Paint by Example: Exemplar-Based Image Editing With Diffusion Models, *Binxin Yang, Shuyang Gu, Bo Zhang, Ting Zhang, Xuejin Chen, Xiaoyan Sun, Dong Chen, Fang Wen*
183. InstructPix2Pix: Learning to Follow Image Editing Instructions, ✧ *Tim Brooks, Aleksander Holynski, Alexei A. Efros*
184. LayoutFormer++: Conditional Graphic Layout Generation via Constraint Serialization and Decoding Space Restriction, *Zhaoyun Jiang, Jiaqi Guo, Shizhao Sun, Huayu Deng, Zhongkai Wu, Vuksan Mijovic, Zijiang James Yang, Jian-Guang Lou, Dongmei Zhang*
185. Self-Guided Diffusion Models, *Vincent Tao Hu, David W. Zhang, Yuki M. Asano, Gertjan J. Burghouts, Cees G. M. Snoek*
186. HOLODIFFUSION: Training a 3D Diffusion Model Using 2D Images, *Animesh Karnewar, Andrea Vedaldi, David Novotny, Niloy J. Mitra*
187. Class-Balancing Diffusion Models, *Yiming Qin, Huangjie Zheng, Jiangchao Yao, Mingyuan Zhou, Ya Zhang*
188. Conditional Image-to-Video Generation With Latent Flow Diffusion Models, *Haomiao Ni, Changhao Shi, Kai Li, Sharon X. Huang, Martin Renqiang Min*
189. Video Probabilistic Diffusion Models in Projected Latent Space, *Sihyun Yu, Kihyuk Sohn, Subin Kim, Jinwoo Shin*
190. Regularized Vector Quantization for Tokenized Image Synthesis, *Jiahui Zhang, Fangneng Zhan, Christian Theobalt, Shijian Lu*
191. EfficientSCI: Densely Connected Network With Space-Time Factorization for Large-Scale Video Snapshot Compressive Imaging, *Lishun Wang, Miao Cao, Xin Yuan*
192. MMVC: Learned Multi-Mode Video Compression With Block-Based Prediction Mode Selection and Density-Adaptive Entropy Coding, *Bowen Liu, Yu Chen, Rakesh Chowdary Machineni, Shiyu Liu, Hun-Seok Kim*
193. Video Compression With Entropy-Constrained Neural Representations, *Carlos Gomes, Roberto Azevedo, Christopher Schroers*
194. WIRE: Wavelet Implicit Neural Representations, *Vishwanath Saragadam, Daniel LeJeune, Jasper Tan, Guha Balakrishnan, Ashok Veeraraghavan, Richard G. Baraniuk*
195. TINC: Tree-Structured Implicit Neural Compression, *Runzhao Yang*
196. CompletionFormer: Depth Completion With Convolutions and Vision Transformers, *Youmin Zhang, Xianda Guo, Matteo Poggi, Zheng Zhu, Guan Huang, Stefano Mattoccia*
197. Lite-Mono: A Lightweight CNN and Transformer Architecture for Self-Supervised Monocular Depth Estimation, *Ning Zhang, Francesco Nex, George Vosselman, Norman Kerle*
198. Global Vision Transformer Pruning With Hessian-Aware Saliency, *Huanrui Yang, Hongxu Yin, Maying Shen, Pavlo Molchanov, Hai Li, Jan Kautz*
199. Lite DETR: An Interleaved Multi-Scale Encoder for Efficient DETR, *Feng Li, Ailing Zeng, Shilong Liu, Hao Zhang, Hongyang Li, Lei Zhang, Lionel M. Ni*
200. PaCa-ViT: Learning Patch-to-Cluster Attention in Vision Transformers, *Ryan Grainger, Thomas Paniagua, Xi Song, Naresh Cuntoor, Mun Wai Lee, Tianfu Wu*
201. Visual Atoms: Pre-Training Vision Transformers With Sinusoidal Waves, *Sora Takashima, Ryo Hayamizu, Nakamasa Inoue, Hirokatsu Kataoka, Rio Yokota*
202. Neuron Structure Modeling for Generalizable Remote Physiological Measurement, *Hao Lu, Zitong Yu, Xuesong Niu, Ying-Cong Chen*
203. Explaining Image Classifiers With Multiscale Directional Image Representation, *Stefan Kolek, Robert Windesheim, Hector Andrade-Loarca, Gitta Kutyniok, Ron Levie*
204. Integrally Pre-Trained Transformer Pyramid Networks, *Yunjie Tian, Lingxi Xie, Zhaozhi Wang, Longhui Wei, Xiaopeng Zhang, Jianbin Jiao, Yaowei Wang, Qi Tian, Qixiang Ye*
205. PartMix: Regularization Strategy to Learn Part Discovery for Visible-Infrared Person Re-Identification, *Minsu Kim, Seungryong Kim, Jungin Park, Seongheon Park, Kwanghoon Sohn*
206. Knowledge Distillation for 6D Pose Estimation by Aligning Distributions of Local Predictions, *Shuxuan Guo, Yinlin Hu, Jose M. Alvarez, Mathieu Salzmann*
207. Focused and Collaborative Feedback Integration for Interactive Image Segmentation, *Qiaoqiao Wei, Hui Zhang, Jun-Hai Yong*
208. PolyFormer: Referring Image Segmentation As Sequential Polygon Generation, *Jiang Liu, Hui Ding, Zhaowei Cai, Yuting Zhang, Ravi Kumar Satzoda, Vijay Mahadevan, R. Manmatha*
209. Devil's on the Edges: Selective Quad Attention for Scene Graph Generation, *Deunsol Jung, Sanghyun Kim, Won Hwa Kim, Minsu Cho*
210. Panoptic Video Scene Graph Generation, *Jingkang Yang, Wenxuan Peng, Xiangtai Li, Zujin Guo, Liangyu Chen, Bo Li, Zheng Ma, Kaiyang Zhou, Wayne Zhang, Chen Change Loy, Ziwei Liu*
211. Generalized Relation Modeling for Transformer Tracking, *Shenyuan Gao, Chunlun Zhou, Jun Zhang*
212. Representation Learning for Visual Object Tracking by Masked Appearance Transfer, *Haojie Zhao, Dong Wang, Huchuan Lu*
213. Unified Mask Embedding and Correspondence Learning for Self-Supervised Video Segmentation, *Liulei Li, Wenguan Wang, Tianfei Zhou, Jianwu Li, Yi Yang*
214. EVAL: Explainable Video Anomaly Localization, *Ashish Singh, Michael J. Jones, Erik G. Learned-Miller*
215. MOSO: Decomposing MOTion, Scene and Object for Video Prediction, *Mingzhen Sun, Weining Wang, Xinxin Zhu, Jing Liu*
216. TarViS: A Unified Approach for Target-Based Video
✧ Segmentation, *Ali Athar, Alexander Hermans, Jonathon Luiten, Deva Ramanan, Bastian Leibe*
217. Efficient Movie Scene Detection Using State-Space Transformers, *Md Mohaiminul Islam, Mahmudul Hasan, Kishan Shamsundar Athrey, Tony Braskich, Gedas Bertasius*
218. Latency Matters: Real-Time Action Forecasting Transformer, ✧ *Harshayu Girase, Nakul Agarwal, Chiho Choi, Karttikeya Mangalam*

219. Temporal Attention Unit: Towards Efficient Spatiotemporal Predictive Learning, *Cheng Tan, Zhangyang Gao, Lirong Wu, Yongjie Xu, Jun Xia, Siyuan Li, Stan Z. Li*
220. Watch or Listen: Robust Audio-Visual Speech Recognition With Visual Corruption Modeling and Reliability Scoring, *Joanna Hong, Minsu Kim, Jeongsoo Choi, Yong Man Ro*
221. ReVISE: Self-Supervised Speech Resynthesis With Visual Input for Universal and Generalized Speech Regeneration, *Wei-Ning Hsu, Tal Remez, Bowen Shi, Jacob Donley, Yossi Adi*
222. SynthVSR: Scaling Up Visual Speech Recognition With Synthetic Supervision, *Xubo Liu, Egor Lakomkin, Konstantinos Vougioukas, Pingchuan Ma, Honglie Chen, Ruiming Xie, Morrie Doulaty, Niko Moritz, Jachym Kolar, Stavros Petridis, Maja Pantic, Christian Fuegen*
223. SVFormer: Semi-Supervised Video Transformer for Action Recognition, *Zhen Xing, Qi Dai, Han Hu, Jingjing Chen, Zuxuan Wu, Yu-Gang Jiang*
224. Collecting Cross-Modal Presence-Absence Evidence for Weakly-Supervised Audio-Visual Event Perception, *Junyu Gao, Mengyuan Chen, Changsheng Xu*
225. Post-Processing Temporal Action Detection, *Sauradip Nag, Xiatian Zhu, Yi-Zhe Song, Tao Xiang*
226. HaLP: Hallucinating Latent Positives for Skeleton-Based Self-Supervised Learning of Actions, *Anshul Shah, Aniket Roy, Ketul Shah, Shlok Mishra, David Jacobs, Anoop Cherian, Rama Chellappa*
227. TriDet: Temporal Action Detection With Relative Boundary Modeling, *Dingfeng Shi, Yujie Zhong, Qiong Cao, Lin Ma, Jia Li, Dacheng Tao*
228. Hybrid Active Learning via Deep Clustering for Video Action Detection, *Aayush J. Rana, Yogesh S. Rawat*
229. Two-Stream Networks for Weakly-Supervised Temporal Action Localization With Semantic-Aware Mechanisms, *Yu Wang, Yadong Li, Hongbin Wang*
230. Weakly Supervised Video Emotion Detection and Prediction via Cross-Modal Temporal Erasing Network, *Zhicheng Zhang, Lijuan Wang, Jufeng Yang*
231. Collaborative Noisy Label Cleaner: Learning Scene-Aware Trailers for Multi-Modal Highlight Detection in Movies, *Bei Gan, Xiujun Shu, Ruizhi Qiao, Haoqian Wu, Keyu Chen, Hanjun Li, Bo Ren*
232. Weakly Supervised Temporal Sentence Grounding With Uncertainty-Guided Self-Training, *Yifei Huang, Lijin Yang, Yoichi Sato*
233. SViTT: Temporal Learning of Sparse Video-Text Transformers, *Yi Li, Kyle Min, Subarna Tripathi, Nuno Vasconcelos*
234. AutoAD: Movie Description in Context, *Tengda Han, Max Bain, ✧ Arsha Nagrani, Gül Varol, Weidi Xie, Andrew Zisserman*
235. Text With Knowledge Graph Augmented Transformer for Video Captioning, *Xin Gu, Guang Chen, Yufei Wang, Libo Zhang, Tiejian Luo, Longyin Wen*
236. StepFormer: Self-Supervised Step Discovery and Localization in Instructional Videos, *Nikita Dvornik, Isma Hadji, Ran Zhang, Konstantinos G. Derpanis, Richard P. Wildes, Allan D. Jepson*
237. Dual Alignment Unsupervised Domain Adaptation for Video-Text Retrieval, *Xiaoshuai Hao, Wanqian Zhang, Dayan Wu, Fei Zhu, Bo Li*
238. Hierarchical Semantic Correspondence Networks for Video Paragraph Grounding, *Chaolei Tan, Zihang Lin, Jian-Fang Hu, Wei-Shi Zheng, Jianhuang Lai*
239. CLIPPING: Distilling CLIP-Based Models With a Student Base for Video-Language Retrieval, *Renjing Pei, Jianzhuang Liu, Weimian Li, Bin Shao, Songcen Xu, Peng Dai, Juwei Lu, Youliang Yan*
240. Learning Emotion Representations From Verbal and Nonverbal Communication, *Sitao Zhang, Yimu Pan, James Z. Wang*
241. Context De-Confounded Emotion Recognition, *Dingkang Yang, Zhaoyu Chen, Yuzheng Wang, Shunli Wang, Mingcheng Li, Siao Liu, Xiao Zhao, Shuai Huang, Zhiyan Dong, Peng Zhai, Lihua Zhang*
242. CiCo: Domain-Aware Sign Language Retrieval via Cross-Lingual Contrastive Learning, *Yiting Cheng, Fangyun Wei, Jianmin Bao, Dong Chen, Wenqiang Zhang*
243. Discovering the Real Association: Multimodal Causal Reasoning in Video Question Answering, *Chuanqi Zang, Hanqing Wang, Mingtao Pei, Wei Liang*
244. LEGO-Net: Learning Regular Rearrangements of Objects in Rooms, *Qihong Anna Wei, Sijie Ding, Jeong Joon Park, Rahul Sajnani, Adrien Poulenard, Srinath Sridhar, Leonidas Guibas*
245. LANA: A Language-Capable Navigator for Instruction Following and Generation, *Xiaohan Wang, Wenguan Wang, Jiayi Shao, Yi Yang*
246. Policy Adaptation From Foundation Model Feedback, *Yuying Ge, Annabella Macaluso, Li Erran Li, Ping Luo, Xiaolong Wang*
247. Token Turing Machines, *Michael S. Ryoo, Keerthana Gopalakrishnan, Kumara Kahatapitiya, Ted Xiao, Kanishka Rao, Austin Stone, Yao Lu, Julian Ibarz, Anurag Arnab*
248. Unicode Analogies: An Anti-Objectivist Visual Reasoning Challenge, *Steven Spratley, Krista A. Ehinger, Tim Miller*
249. Exploring the Effect of Primitives for Compositional Generalization in Vision-and-Language, *Chuanhao Li, Zhen Li, Chenchen Jing, Yunde Jia, Yuwei Wu*
250. VQACL: A Novel Visual Question Answering Continual Learning Setting, *Xi Zhang, Feifei Zhang, Changsheng Xu*
251. MaPLE: Multi-Modal Prompt Learning, *Muhammad Uzair Khattak, Hanoona Rasheed, Muhammad Maaz, Salman Khan, Fahad Shahbaz Khan*
252. Meta-Personalizing Vision-Language Models to Find Named Instances in Video, *Chun-Hsiao Yeh, Bryan Russell, Josef Sivic, Fabian Caba Heilbron, Simon Jenni*
253. Understanding and Improving Visual Prompting: A Label-Mapping Perspective, *Aochuan Chen, Yuguang Yao, Pin-Yu Chen, Yihua Zhang, Sijia Liu*
254. RefTeacher: A Strong Baseline for Semi-Supervised Referring Expression Comprehension, *Jiamu Sun, Gen Luo, Yiyi Zhou, Xiaoshuai Sun, Guannan Jiang, Zhiyu Wang, Rongrong Ji*
255. Leveraging per Image-Token Consistency for Vision-Language Pre-Training, *Yunhao Gou, Tom Ko, Hansi Yang, James Kwok, Yu Zhang, Mingxuan Wang*
256. Improving Visual Grounding by Encouraging Consistent Gradient-Based Explanations, *Ziyan Yang, Kushal Kafle, Franck Dernoncourt, Vicente Ordonez*
257. Image as a Foreign Language: BEiT Pretraining for Vision and Vision-Language Tasks, *Wenhui Wang, Hangbo Bao, Li Dong, Johan Bjorck, Zhiliang Peng, Qiang Liu, Kriti Aggarwal, Owais Khan Mohammed, Saksham Singhal, Subhojit Som, Furu Wei*

258. Language in a Bottle: Language Model Guided Concept Bottlenecks for Interpretable Image Classification, *Yue Yang, Artemis Panagopoulou, Shenghao Zhou, Daniel Jin, Chris Callison-Burch, Mark Yatskar*
259. Shepherding Slots to Objects: Towards Stable and Robust Object-Centric Learning, *Jinwoo Kim, Janghyuk Choi, Ho-Jin Choi, Seon Joo Kim*
260. Learning Visual Representations via Language-Guided Sampling, *Mohamed El Banani, Karan Desai, Justin Johnson*
261. L-Colns: Language-Based Colorization With Instance Awareness, *Zheng Chang, Shuchen Weng, Peixuan Zhang, Yu Li, Si Li, Boxin Shi*
262. EDA: Explicit Text-Decoupling and Dense Alignment for 3D Visual Grounding, *Yanmin Wu, Xinhua Cheng, Renrui Zhang, Zesen Cheng, Jian Zhang*
263. MSINet: Twins Contrastive Search of Multi-Scale Interaction for Object ReID, *Jianyang Gu, Kai Wang, Hao Luo, Chen Chen, Wei Jiang, Yuqiang Fang, Shanghang Zhang, Yang You, Jian Zhao*
264. Unifying Vision, Text, and Layout for Universal Document Processing, *Zineng Tang, Ziyi Yang, Guoxin Wang, Yuwei Fang, Yang Liu, Chenguang Zhu, Michael Zeng, Cha Zhang, Mohit Bansal*
265. RA-CLIP: Retrieval Augmented Contrastive Language-Image Pre-Training, *Chen-Wei Xie, Siyang Sun, Xiong Xiong, Yun Zheng, Deli Zhao, Jingren Zhou*
266. Fine-Grained Image-Text Matching by Cross-Modal Hard Aligning Network, *Zhengxin Pan, Fangyu Wu, Bailing Zhang*
267. Text-Guided Unsupervised Latent Transformation for Multi-Attribute Image Manipulation, *Xiwen Wei, Zhen Xu, Cheng Liu, Si Wu, Zhiwen Yu, Hau San Wong*
268. Improving Image Recognition by Retrieving From Web-Scale Image-Text Data, *Ahmet Iscen, Alireza Fathi, Cordelia Schmid*
269. Pic2Word: Mapping Pictures to Words for Zero-Shot Composed Image Retrieval, *Kuniaki Saito, Kihyuk Sohn, Xiang Zhang, Chun-Liang Li, Chen-Yu Lee, Kate Saenko, Tomas Pfister*
270. DATE: Domain Adaptive Product Seeker for E-Commerce, *Haoyuan Li, Hao Jiang, Tao Jin, Mengyan Li, Yan Chen, Zhijie Lin, Yang Zhao, Zhou Zhao*
271. Multimodality Helps Unimodality: Cross-Modal Few-Shot Learning With Multimodal Models, *Zhiqiu Lin, Samuel Yu, Zhiyi Kuang, Deepak Pathak, Deva Ramanan*
272. Finetune Like You Pretrain: Improved Finetuning of Zero-Shot Vision Models, *Sachin Goyal, Ananya Kumar, Sankalp Garg, Zico Kolter, Aditi Raghunathan*
273. DeepSolo: Let Transformer Decoder With Explicit Points Solo for Text Spotting, *Maoyuan Ye, Jing Zhang, Shanshan Zhao, Juhua Liu, Tongliang Liu, Bo Du, Dacheng Tao*
274. EVA: Exploring the Limits of Masked Visual Representation Learning at Scale, *Yuxin Fang, Wen Wang, Binhui Xie, Quan Sun, Ledell Wu, Xinggang Wang, Tiejun Huang, Xinlong Wang, Yue Cao*
275. R²Former: Unified Retrieval and Reranking Transformer for Place Recognition, *Sijie Zhu, Linjie Yang, Chen Chen, Mubarak Shah, Xiaohui Shen, Heng Wang*
276. Open-Set Fine-Grained Retrieval via Prompting Vision-Language Evaluator, *Shijie Wang, Jianlong Chang, Haojie Li, Zhihui Wang, Wanli Ouyang, Qi Tian*
277. Open-Category Human-Object Interaction Pre-Training via Language Modeling Framework, *Sipeng Zheng, Boshen Xu, Qin Jin*
278. Neural Congealing: Aligning Images to a Joint Semantic Atlas, *Dolev Ofri-Amar, Michal Geyer, Yoni Kasten, Tali Dekel*
279. Open Vocabulary Semantic Segmentation With Patch Aligned Contrastive Learning, *Jishnu Mukhoti, Tsung-Yu Lin, Omid Poursaeed, Rui Wang, Ashish Shah, Philip H.S. Torr, Ser-Nam Lim*
280. Semantic Human Parsing via Scalable Semantic Transfer Over Multiple Label Domains, *Jie Yang, Chaoqun Wang, Zhen Li, Junle Wang, Ruimao Zhang*
281. Explicit Visual Prompting for Low-Level Structure Segmentations, *Weihuang Liu, Xi Shen, Chi-Man Pun, Xiaodong Cun*
282. FreeSeg: Unified, Universal and Open-Vocabulary Image Segmentation, *Jie Qin, Jie Wu, Pengxiang Yan, Ming Li, Ren Yuxi, Xuefeng Xiao, Yitong Wang, Rui Wang, Shilei Wen, Xin Pan, Xingang Wang*
283. Zero-Shot Referring Image Segmentation With Global-Local Context Features, *Seonghoon Yu, Paul Hongsuck Seo, Jeany Son*
284. DeJaVu: Conditional Regenerative Learning to Enhance Dense Prediction, *Shubhankar Borse, Debasmit Das, Hyojin Park, Hong Cai, Risheek Garrepalli, Fatih Porikli*
285. Meta Compositional Referring Expression Segmentation, *Li Xu, Mark He Huang, Xindi Shang, Zehuan Yuan, Ying Sun, Jun Liu*
286. Interactive Segmentation As Gaussian Process Classification, *Minghao Zhou, Hong Wang, Qian Zhao, Yuexiang Li, Yawen Huang, Deyu Meng, Yefeng Zheng*
287. Semantic-Promoted Debiasing and Background Disambiguation for Zero-Shot Instance Segmentation, *Shuting He, Henghui Ding, Wei Jiang*
288. Principles of Forgetting in Domain-Incremental Semantic Segmentation in Adverse Weather Conditions, *Tobias Kalb, Jürgen Beyerer*
289. AttentionShift: Iteratively Estimated Part-Based Attention Map for Pointly Supervised Instance Segmentation, *Mingxiang Liao, Zonghao Guo, Yuze Wang, Peng Yuan, Bailan Feng, Fang Wan*
290. PIDNet: A Real-Time Semantic Segmentation Network Inspired by PID Controllers, *Jiacong Xu, Zixiang Xiong, Shankar P. Bhattacharyya*
291. Leveraging Hidden Positives for Unsupervised Semantic Segmentation, *Hyun Seok Seong, WonJun Moon, SuBeen Lee, Jae-Pil Heo*
292. Understanding Imbalanced Semantic Segmentation Through Neural Collapse, *Zhisheng Zhong, Jiequan Cui, Yibo Yang, Xiaoyang Wu, Xiaojuan Qi, Xiangyu Zhang, Jiaya Jia*
293. Balancing Logit Variation for Long-Tailed Semantic Segmentation, *Yuchao Wang, Jingjing Fei, Haochen Wang, Wei Li, Tianpeng Bao, Liwei Wu, Rui Zhao, Yujun Shen*
294. Boundary-Enhanced Co-Training for Weakly Supervised Semantic Segmentation, *Shenghai Rong, Bohai Tu, Zilei Wang, Junjie Li*
295. Conflict-Based Cross-View Consistency for Semi-Supervised Semantic Segmentation, *Zicheng Wang, Zhen Zhao, Xiaoxia Xing, Dong Xu, Xiangyu Kong, Luping Zhou*
296. Learning Multi-Modal Class-Specific Tokens for Weakly Supervised Dense Object Localization, *Lian Xu, Wanli Ouyang, Mohammed Bennamoun, Farid Boussaid, Dan Xu*

297. WinCLIP: Zero-/Few-Shot Anomaly Classification and Segmentation, *Jongheon Jeong, Yang Zou, Taewan Kim, Dongqing Zhang, Avinash Ravichandran, Onkar Dabeer*
298. DualRel: Semi-Supervised Mitochondria Segmentation From a Prototype Perspective, *Huayu Mai, Rui Sun, Tianzhu Zhang, Zhiwei Xiong, Feng Wu*
299. Distilling Self-Supervised Vision Transformers for Weakly-Supervised Few-Shot Classification & Segmentation, *Dahyun Kang, Piotr Koniusz, Minsu Cho, Naila Murray*
300. Co-Salient Object Detection With Uncertainty-Aware Group Exchange-Masking, *Yang Wu, Huihui Song, Bo Liu, Kaihua Zhang, Dong Liu*
301. Supervised Masked Knowledge Distillation for Few-Shot Transformers, *Han Lin, Guangxing Han, Jiawei Ma, Shiyuan Huang, Xudong Lin, Shih-Fu Chang*
302. Modeling the Distributional Uncertainty for Salient Object Detection Models, *Xinyu Tian, Jing Zhang, Mochu Xiang, Yuchao Dai*
303. Weak-Shot Object Detection Through Mutual Knowledge Transfer, *Xuanyi Du, Weitao Wan, Chong Sun, Chen Li*
304. CAT: LoCalization and IdentificAtion Cascade Detection Transformer for Open-World Object Detection, *Shuailei Ma, Yuefeng Wang, Ying Wei, Jiaqi Fan, Thomas H. Li, Hongli Liu, Fanbing Lv*
305. Adaptive Sparse Pairwise Loss for Object Re-Identification, *Xiao Zhou, Yujie Zhong, Zhen Cheng, Fan Liang, Lin Ma*
306. DETRs With Hybrid Matching, *Ding Jia, Yuhui Yuan, Haodi He, Xiaopei Wu, Haojun Yu, Weihong Lin, Lei Sun, Chao Zhang, Han Hu*
307. Generating Features With Increased Crop-Related Diversity for Few-Shot Object Detection, *Jingyi Xu, Hieu Le, Dimitris Samaras*
308. ScaleKD: Distilling Scale-Aware Knowledge in Small Object Detector, *Yichen Zhu, Qiqi Zhou, Ning Liu, Zhiyuan Xu, Zhicai Ou, Xiaofeng Mou, Jian Tang*
309. Multiclass Confidence and Localization Calibration for Object Detection, *Bimsara Pathiraja, Malitha Gunawardhana, Muhammad Haris Khan*
310. Open-Set Representation Learning Through Combinatorial Embedding, *Geeho Kim, Junoh Kang, Bohyung Han*
311. ProD: Prompting-To-Disentangle Domain Knowledge for Cross-Domain Few-Shot Image Classification, *Tianyi Ma, Yifan Sun, Zongxin Yang, Yi Yang*
312. Visual Language Pretrained Multiple Instance Zero-Shot Transfer for Histopathology Images, *Ming Y. Lu, Bowen Chen, Andrew Zhang, Drew F. K. Williamson, Richard J. Chen, Tong Ding, Long Phi Le, Yung-Sung Chuang, Faisal Mahmood*
313. FFF: Fragment-Guided Flexible Fitting for Building Complete Protein Structures, *Weijie Chen, Xinyan Wang, Yuhang Wang*
314. Pseudo-Label Guided Contrastive Learning for Semi-Supervised Medical Image Segmentation, *Hritam Basak, Zhaozheng Yin*
315. Hierarchical Discriminative Learning Improves Visual
✧ Representations of Biomedical Microscopy, *Cheng Jiang, Xinhai Hou, Akhil Kondepudi, Asadur Chowdury, Christian W. Freudiger, Daniel A. Orringer, Honglak Lee, Todd C. Hollon*
316. KiUT: Knowledge-Injected U-Transformer for Radiology Report Generation, *Zhongzhen Huang, Xiaofan Zhang, Shaoting Zhang*
317. Image Quality-Aware Diagnosis via Meta-Knowledge Co-Embedding, *Haoxuan Che, Siyu Chen, Hao Chen*
318. Interventional Bag Multi-Instance Learning on Whole-Slide
✧ Pathological Images, *Tiancheng Lin, Zhimiao Yu, Hongyu Hu, Yi Xu, Chang-Wen Chen*
319. Visual Prompt Tuning for Generative Transfer Learning, *Kihyuk Sohn, Huiwen Chang, José Lezama, Luisa Polania, Han Zhang, Yuan Hao, Irfan Essa, Lu Jiang*
320. LiNe: Out-of-Distribution Detection by Leveraging Important Neurons, *Yong Hyun Ahn, Gyeong-Moon Park, Seong Tae Kim*
321. GCFagg: Global and Cross-View Feature Aggregation for Multi-View Clustering, *Weiqing Yan, Yuanyang Zhang, Chenlei Lv, Chang Tang, Guanghui Yue, Liang Liao, Weisi Lin*
322. Exploring and Exploiting Uncertainty for Incomplete Multi-View Classification, *Mengyao Xie, Zongbo Han, Changqing Zhang, Yichen Bai, Qinghua Hu*
323. BiCro: Noisy Correspondence Rectification for Multi-Modality Data via Bi-Directional Cross-Modal Similarity Consistency, *Shuo Yang, Zhaopan Xu, Kai Wang, Yang You, Hongxun Yao, Tongliang Liu, Min Xu*
324. Bi-Directional Distribution Alignment for Transductive Zero-Shot Learning, *Zhicai Wang, Yanbin Hao, Tingting Mu, Ouxiang Li, Shuo Wang, Xiangnan He*
325. HIER: Metric Learning Beyond Class Labels via Hierarchical Regularization, *Sungyeon Kim, Boseung Jeong, Suha Kwak*
326. MaskCon: Masked Contrastive Learning for Coarse-Labelled Dataset, *Chen Feng, Ioannis Patras*
327. Class Prototypes Based Contrastive Learning for Classifying Multi-Label and Fine-Grained Educational Videos, *Rohit Gupta, Anirban Roy, Claire Christensen, Sujeong Kim, Sarah Gerard, Madeline Cincebeaux, Ajay Divakaran, Todd Grindal, Mubarak Shah*
328. Learning From Noisy Labels With Decoupled Meta Label Purifier, *Yuanpeng Tu, Boshen Zhang, Yuxi Li, Liang Liu, Jian Li, Yabiao Wang, Chengjie Wang, Cai Rong Zhao*
329. SuperDisco: Super-Class Discovery Improves Visual Recognition for the Long-Tail, *Yingjun Du, Jiayi Shen, Xiantong Zhen, Cees G. M. Snoek*
330. Why Is the Winner the Best? *Matthias Eisenmann, Annika Reinke, Vivienn Weru, Minu D. Tizabi, Fabian Isensee, Tim J. Adler, Sharib Ali, Vincent Andrearczyk, Marc Aubreville, Ujjwal Baid, Spyridon Bakas, Niranjana Balu, Sophia Bano, Jorge Bernal, Sebastian Bodenstedt, Alessandro Casella, Veronika Cheplygina, Marie Daum, Marleen de Bruijne, Adrien Depeursinge, Reuben Dorent, Jan Egger, David G. Ellis, Sandy Engelhardt, Melanie Ganz, Noha Ghatwary, Gabriel Girard, Patrick Godau, Anubha Gupta, Lasse Hansen, Kanako Harada, Mattias P. Heinrich, Nicholas Heller, Alessa Hering, Arnaud Huault, Pierre Jannin, Ali Emre Kavur, Oldřich Kodým, Michal Kozubek, Jianing Li, Hongwei Li, Jun Ma, Carlos Martín-Isla, Bjoern Menze, Alison Noble, Valentin Oreiller, Nicolas Padoy, Sarthak Pati, Kelly Payette, Tim Radsch, Jonathan Rafael-Patiño, Vivek Singh Bawa, Stefanie Speidel, Carole H. Sudre, Kimberlin van Wijnjen, Martin Wagner, Donglai Wei, Amine Yamlahi, Moi Hoon Yap, Chun Yuan, Maximilian Zenk, Aneeq Zia, David Zimmerer, Dogu Baran Aydogan, Binod Bhattarai, Louise Bloch, Raphael Brüngel, Jihoon Cho, Chanyeol Choi, Qi Dou, Ivan Ezhov, Christoph M. Friedrich, Clifton D. Fuller, Rebati Raman Gaire, Adrian Galdan, Álvaro García Faura, Maria Grammatikopoulou, Seulgi Hong, Mostafa Jahanifar, Ikbeom Jang, Abdolrahim Kadkhodamohammadi, Inha Kang, Florian Kofler, Satoshi Kondo, Hugo Kuijff, Mingxing Li, Minh Luu, Tomaz Martinčič, Pedro Morais, Mohamed A. Naser,*

- Bruno Oliveira, David Owen, Subeen Pang, Jinah Park, Sung-Hong Park, Szymon Plotka, Elodie Puybareau, Nasir Rajpoot, Kanghyun Ryu, Numan Saeed, Adam Shephard, Pengcheng Shi, Dejan Štepec, Ronast Subedi, Guillaume Tochon, Helena R. Torres, Helene Urien, João L. Vilaça, Kareem A. Wahid, Haojie Wang, Jiacheng Wang, Liansheng Wang, Xiyue Wang, Benedikt Wiestler, Marek Wodzinski, Fangfang Xia, Juanying Xie, Zhiwei Xiong, Sen Yang, Yanwu Yang, Zixuan Zhao, Klaus Maier-Hein, Paul F. Jäger, Annette Kopp-Schneider, Lena Maier-Hein
331. Balanced Product of Calibrated Experts for Long-Tailed Recognition, Emanuel Sanchez Aimar, Arvi Jonnarth, Michael Felsberg, Marco Kuhlmann
332. Transfer Knowledge From Head to Tail: Uncertainty Calibration Under Long-Tailed Distribution, Jiahao Chen, Bing Su
333. FREDOM: Fairness Domain Adaptation Approach to Semantic Scene Understanding, Thanh-Dat Truong, Ngan Le, Bhiksha Raj, Jackson Cothren, Khoa Luu
334. COT: Unsupervised Domain Adaptation With Clustering and Optimal Transport, Yang Liu, Zhipeng Zhou, Baigui Sun
335. MHPL: Minimum Happy Points Learning for Active Source Free Domain Adaptation, Fan Wang, Zhongyi Han, Zhiyan Zhang, Rundong He, Yilong Yin
336. Upcycling Models Under Domain and Category Shift, Sanqing Qu, Tianpei Zou, Florian Röhrbein, Cewu Lu, Guang Chen, Dacheng Tao, Changjun Jiang
337. PMR: Prototypical Modal Rebalance for Multimodal Learning, Yunfeng Fan, Wenchao Xu, Haozhao Wang, Junxiao Wang, Song Guo
338. MMANet: Margin-Aware Distillation and Modality-Aware Regularization for Incomplete Multimodal Learning, Shicai Wei, Chunbo Luo, Yang Luo
339. Feature Alignment and Uniformity for Test Time Adaptation, Shuai Wang, Daoan Zhang, Zipei Yan, Jianguo Zhang, Rui Li
340. Revisiting Prototypical Network for Cross Domain Few-Shot Learning, Fei Zhou, Peng Wang, Lei Zhang, Wei Wei, Yanning Zhang
341. A Whac-a-Mole Dilemma: Shortcuts Come in Multiples Where Mitigating One Amplifies Others, Zhiheng Li, Ivan Evtimov, Albert Gordo, Caner Hazirbas, Tal Hassner, Cristian Canton Ferrer, Chenliang Xu, Mark Ibrahim
342. Independent Component Alignment for Multi-Task Learning, Dmitry Senushkin, Nikolay Patakin, Arseny Kuznetsov, Anton Konushin
343. MDL-NAS: A Joint Multi-Domain Learning Framework for Vision Transformer, Shiguang Wang, Tao Xie, Jian Cheng, Xingcheng Zhang, Haijun Liu
344. MELTR: Meta Loss Transformer for Learning to Fine-Tune Video Foundation Models, Dohwan Ko, Joonmyung Choi, Hyeong Kyu Choi, Kyoung-Woon On, Byungseok Roh, Hyunwoo J. Kim
345. 1% VS 100%: Parameter-Efficient Low Rank Adapter for Dense Predictions, Dongshuo Yin, Yiran Yang, Zhechao Wang, Hongfeng Yu, Kaiwen Wei, Xian Sun
346. Rebalancing Batch Normalization for Exemplar-Based Class-Incremental Learning, Sungmin Cha, Sungjun Cho, Dasol Hwang, Sunwon Hong, Moontae Lee, Taesup Moon
347. Partial Network Cloning, Jingwen Ye, Songhua Liu, Xinchao Wang
348. ERM-KTP: Knowledge-Level Machine Unlearning via Knowledge Transfer, Shen Lin, Xiaoyu Zhang, Chenyang Chen, Xiaofeng Chen, Willy Susilo
349. Rethinking Feature-Based Knowledge Distillation for Face Recognition, Jingzhi Li, Zidong Guo, Hui Li, Seungju Han, Ji-won Baek, Min Yang, Ran Yang, Sungjoo Suh
350. Regularizing Second-Order Influences for Continual Learning, Zhicheng Sun, Yadong Mu, Gang Hua
351. Generalization Matters: Loss Minima Flattening via Parameter Hybridization for Efficient Online Knowledge Distillation, Tianli Zhang, Mengqi Xue, Jiangtao Zhang, Haoqi Zhang, Yu Wang, Lechao Cheng, Jie Song, Mingli Song
352. Decoupling Learning and Remembering: A Bilevel Memory Framework With Knowledge Projection for Task-Incremental Learning, Wenju Sun, Qingyong Li, Jing Zhang, Wen Wang, Yangli-ao Geng
353. On the Stability-Plasticity Dilemma of Class-Incremental Learning, Dongwan Kim, Bohyung Han
354. Simulated Annealing in Early Layers Leads to Better Generalization, Amir M. Sarfi, Zahra Karimpour, Muawiz Chaudhary, Nasir M. Khalid, Mirco Ravanelli, Sudhir Mudur, Eugene Belilovsky
355. Frustratingly Easy Regularization on Representation Can Boost Deep Reinforcement Learning, Qiang He, Huangyuan Su, Jieyu Zhang, Xinwen Hou
356. Tunable Convolutions With Parametric Multi-Loss Optimization, Matteo Maggioni, Thomas Tanay, Francesca Babiloni, Steven McDonagh, Aleš Leonardis
357. Re-Basin via Implicit Sinkhorn Differentiation, Fidel A. Guerrero Peña, Heitor Rapela Medeiros, Thomas Dubail, Masih Aminbeidokhti, Eric Granger, Marco Pedersoli
358. Gradient Norm Aware Minimization Seeks First-Order Flatness and Improves Generalization, Xingxuan Zhang, Renzhe Xu, Han Yu, Hao Zou, Peng Cui
359. AstroNet: When Astrocyte Meets Artificial Neural Network, Mengqiao Han, Liyuan Pan, Xiabi Liu
360. Network Expansion for Practical Training Acceleration, Ning Ding, Yehui Tang, Kai Han, Chao Xu, Yunhe Wang
361. Defining and Quantifying the Emergence of Sparse Concepts in DNNs, Jie Ren, Mingjie Li, Qirui Chen, Huiqi Deng, Quanshi Zhang
362. Samples With Low Loss Curvature Improve Data Efficiency, Isha Garg, Kaushik Roy
363. Masked Images Are Counterfactual Samples for Robust Fine-Tuning, Yao Xiao, Ziyi Tang, Pengxu Wei, Cong Liu, Liang Lin
364. Bias Mimicking: A Simple Sampling Approach for Bias Mitigation, Maan Qraitem, Kate Saenko, Bryan A. Plummer
365. NoisyQuant: Noisy Bias-Enhanced Post-Training Activation Quantization for Vision Transformers, Yijiang Liu, Huanrui Yang, Zhen Dong, Kurt Keutzer, Li Du, Shanghang Zhang
366. Practical Network Acceleration With Tiny Sets, Guo-Hua Wang, Jianxin Wu
367. TeSLA: Test-Time Self-Learning With Automatic Adversarial Augmentation, Devavrat Tomar, Guillaume Vray, Behzad Bozorgtabar, Jean-Philippe Thiran
368. Discriminator-Cooperated Feature Map Distillation for GAN Compression, Tie Hu, Mingbao Lin, Lizhou You, Fei Chao, Rongrong Ji

389. Cooperation or Competition: Avoiding Player Domination for Multi-Target Robustness via Adaptive Budgets, *Yimu Wang, Dinghuai Zhang, Yihan Wu, Heng Huang, Hongyang Zhang*
390. Discrete Point-Wise Attack Is Not Enough: Generalized Manifold Adversarial Attack for Face Recognition, *Qian Li, Yuxiao Hu, Ye Liu, Dongxiao Zhang, Xin Jin, Yuntian Chen*
391. RIATIG: Reliable and Imperceptible Adversarial Text-to-Image Generation With Natural Prompts, *Han Liu, Yuhao Wu, Shixuan Zhai, Bo Yuan, Ning Zhang*
392. CLIP2Protect: Protecting Facial Privacy Using Text-Guided Makeup via Adversarial Latent Search, *Fahad Shamshad, Muzammal Naseer, Karthik Nandakumar*
393. TruFor: Leveraging All-Round Clues for Trustworthy Image Forgery Detection and Localization, *Fabrizio Guillaro, Davide Cozzolino, Avneesh Sud, Nicholas Dufour, Luisa Verdoliva*

Notes:

[illegible]

1230–1400 Poster Switch/Setup (West Exhibit Hall)

1400–1600 Demos (West Exhibit Hall Demo Area)

- Same as morning demos (see page 48).

1400–1500 Plenary-Thu-PM (East Exhibit Halls A–B)

Panel: Scientific Discovery and the Environment

Moderator: Vladlen Koltun (*Apple*)

Participants:

- Elizabeth A. Barnes (*Colorado State Univ.*)
- Sara Beery (*MIT*)
- Josh Bloom (*Univ. of California, Berkeley*)
- Kyle Cranmer (*Univ. of Wisconsin-Madison*)

1500–1600 Oral-Award-Thu (East Exhibit Halls A–B)

Award candidate paper presentations

Format: 10 min. presentation (including questions)

1. [1500] DreamBooth: Fine Tuning Text-to-Image Diffusion Models for Subject-Driven Generation, *Nataniel Ruiz, Yuanzhen Li, Varun Jampani, Yael Pritch, Michael Rubinstein, Kfir Aberman*
2. [1510] On Distillation of Guided Diffusion Models, *Chenlin Meng, Robin Rombach, Ruiqi Gao, Diederik Kingma, Stefano Ermon, Jonathan Ho, Tim Salimans*
3. [1520] Visual Programming: Compositional Visual Reasoning Without Training, *Tanmay Gupta, Aniruddha Kembhavi*
4. [1530] What Can Human Sketches Do for Object Detection? *Pinaki Nath Chowdhury, Ayan Kumar Bhunia, Aneeshan Sain, Subhadeep Koley, Tao Xiang, Yi-Zhe Song*
5. [1540] Data-Driven Feature Tracking for Event Cameras, *Nico Messikommer, Carter Fang, Mathias Gehrig, Davide Scaramuzza*
6. [1550] Integral Neural Networks, *Kirill Solodskikh, Azim Kurbanov, Ruslan Aydarkhanov, Irina Zhelavskaya, Yury Parfenov, Dehua Song, Stamatios Lefkimmiatis*

1600–1630 Afternoon Break (West Exhibit Hall)

1630–1830 Poster-Thu-PM (West Exhibit Hall)

✧ - Highlight paper (check it out)

🏆 - Award candidate paper (see award sessions)

1. High-Fidelity Event-Radiance Recovery via Transient Event Frequency, *Jin Han, Yuta Asano, Boxin Shi, Yinqiang Zheng, Imari Sato*
2. RobustNeRF: Ignoring Distractors With Robust Losses, *Sara ✧ Sabour, Suhani Vora, Daniel Duckworth, Ivan Krasin, David J. Fleet, Andrea Tagliasacchi*
3. NeRDi: Single-View NeRF Synthesis With Language-Guided Diffusion As General Image Priors, *Congyue Deng, Chiyu "Max" Jiang, Charles R. Qi, Xinchun Yan, Yin Zhou, Leonidas Guibas, Dragomir Anguelov*
4. GM-NeRF: Learning Generalizable Model-Based Neural Radiance Fields From Multi-View Images, *Jianchuan Chen, Wentao Yi, Liqian Ma, Xu Jia, Huchuan Lu*
5. MixNeRF: Modeling a Ray With Mixture Density for Novel View Synthesis From Sparse Inputs, *Seunghyeon Seo, Donghoon Han, Yeonjin Chang, Nojun Kwak*

6. SPIn-NeRF: Multiview Segmentation and Perceptual Inpainting With Neural Radiance Fields, *Ashkan Mirzaei, Tristan Aumentado-Armstrong, Konstantinos G. Derpanis, Jonathan Kelly, Marcus A. Brubaker, Igor Gilitschenski, Alex Levinshtein*
7. Masked Wavelet Representation for Compact Neural Radiance Fields, *Daniel Rho, Byeonghyeon Lee, Seungtae Nam, Joo Chan Lee, Jong Hwan Ko, Eunbyung Park*
8. PaletteNeRF: Palette-Based Appearance Editing of Neural Radiance Fields, *Zhengfei Kuang, Fujun Luan, Sai Bi, Zhixin Shu, Gordon Wetzstein, Kalyan Sunkavalli*
9. SteerNeRF: Accelerating NeRF Rendering via Smooth Viewpoint Trajectory, *Sicheng Li, Hao Li, Yue Wang, Yiyi Liao, Lu Yu*
10. Transforming Radiance Field With Lipschitz Network for ✧ Photorealistic 3D Scene Stylization, *Zicheng Zhang, Yinglu Liu, Congying Han, Yingwei Pan, Tiande Guo, Ting Yao*
11. Occlusion-Free Scene Recovery via Neural Radiance Fields, *Chengxuan Zhu, Renjie Wan, Yunkai Tang, Boxin Shi*
12. TriVol: Point Cloud Rendering via Triple Volumes, *Tao Hu, Xiaogang Xu, Ruihang Chu, Jiaya Jia*
13. DyNCA: Real-Time Dynamic Texture Synthesis Using Neural Cellular Automata, *Ehsan Pajouheshgar, Yitao Xu, Tong Zhang, Sabine Süsstrunk*
14. Neural Scene Chronology, *Haotong Lin, Qianqian Wang, Ruojin Cai, Sida Peng, Hadar Averbuch-Elor, Xiaowei Zhou, Noah Snively*
15. ReLight My NeRF: A Dataset for Novel View Synthesis and ✧ Relighting of Real World Objects, *Marco Toschi, Riccardo De Matteo, Riccardo Spezialetti, Daniele De Gregorio, Luigi Di Stefano, Samuele Salti*
16. ORCa: Glossy Objects As Radiance-Field Cameras, *Kushagra Tiwary, Akshat Dave, Nikhil Behari, Tzofi Klinghoffer, Ashok Veeraraghavan, Ramesh Raskar*
17. Nighttime Smartphone Reflective Flare Removal Using Optical ✧ Center Symmetry Prior, *Yuekun Dai, Yihang Luo, Shangchen Zhou, Chongyi Li, Chen Change Loy*
18. SunStage: Portrait Reconstruction and Relighting Using the Sun as a Light Stage, *Yifan Wang, Aleksander Holynski, Xiuming Zhang, Xuaner Zhang*
19. The Differentiable Lens: Compound Lens Search Over Glass Surfaces and Materials for Object Detection, *Geoffroi Côté, Fahim Mannan, Simon Thibault, Jean-François Lalonde, Felix Heide*
20. Teleidoscopic Imaging System for Microscale 3D Shape Reconstruction, *Ryo Kawahara, Meng-Yu Jennifer Kuo, Shohei Nobuhara*
21. Looking Through the Glass: Neural Surface Reconstruction Against High Specular Reflections, *Jiaxiong Qiu, Peng-Tao Jiang, Yifan Zhu, Ze-Xin Yin, Ming-Ming Cheng, Bo Ren*
22. NeuralUDF: Learning Unsigned Distance Fields for Multi-View Reconstruction of Surfaces With Arbitrary Topologies, *Xiaoxiao Long, Cheng Lin, Lingjie Liu, Yuan Liu, Peng Wang, Christian Theobalt, Taku Komura, Wenping Wang*
23. Sphere-Guided Training of Neural Implicit Surfaces, *Andreea Dogaru, Andrei-Timotei Ardelean, Savva Ignatyev, Egor Zakharov, Evgeny Burnaev*
24. OReX: Object Reconstruction From Planar Cross-Sections Using Neural Fields, *Haim Sawdayee, Amir Vaxman, Amit H. Bermano*

25. Persistent Nature: A Generative Model of Unbounded 3D Worlds, *Lucy Chai, Richard Tucker, Zhengqi Li, Phillip Isola, Noah Snavely*
26. 3D Neural Field Generation Using Triplane Diffusion, *J. Ryan Shue, Eric Ryan Chan, Ryan Po, Zachary Ankner, Jiajun Wu, Gordon Wetzstein*
27. Diffusion-Based Signed Distance Fields for 3D Shape Generation, *Jaehyeok Shim, Changwoo Kang, Kyungdon Joo*
28. Efficient View Synthesis and 3D-Based Multi-Frame Denoising With Multiplane Feature Representations, *Thomas Tanay, Aleš Leonardis, Matteo Maggioni*
29. Dream3D: Zero-Shot Text-to-3D Synthesis Using 3D Shape Prior and Text-to-Image Diffusion Models, *Jiale Xu, Xintao Wang, Weihao Cheng, Yan-Pei Cao, Ying Shan, Xiaohu Qie, Shenghua Gao*
30. SINE: Semantic-Driven Image-Based NeRF Editing With Prior-Guided Editing Field, *Chong Bao, Yinda Zhang, Bangbang Yang, Tianxing Fan, Zesong Yang, Hujun Bao, Guofeng Zhang, Zhaopeng Cui*
31. 3D Highlighter: Localizing Regions on 3D Shapes via Text ✨ Descriptions, *Dale Decatur, Itai Lang, Rana Hanocka*
32. Self-Supervised Geometry-Aware Encoder for Style-Based 3D GAN Inversion, *Yushi Lan, Xuyi Meng, Shuai Yang, Chen Change Loy, Bo Dai*
33. PanoHead: Geometry-Aware 3D Full-Head Synthesis in 360°, *Sizhe An, Hongyi Xu, Yichun Shi, Guoxian Song, Umit Y. Ogras, Linjie Luo*
34. StyleGene: Crossover and Mutation of Region-Level Facial ✨ Genes for Kinship Face Synthesis, *Hao Li, Xianxu Hou, Zepeng Huang, Linlin Shen*
35. Parameter Efficient Local Implicit Image Function Network for Face Segmentation, *Mausoom Sarkar, Nikitha SR, Mayur Hemani, Rishabh Jain, Balaji Krishnamurthy*
36. Graphics Capsule: Learning Hierarchical 3D Face Representations From 2D Images, *Chang Yu, Xiangyu Zhu, Xiaomei Zhang, Zhaoxiang Zhang, Zhen Lei*
37. Next3D: Generative Neural Texture Rasterization for 3D-Aware ✨ Head Avatars, *Jingxiang Sun, Xuan Wang, Lizhen Wang, Xiaoyu Li, Yong Zhang, Hongwen Zhang, Yebin Liu*
38. Learning Neural Parametric Head Models, *Simon Giebenhain, Tobias Kirschstein, Markos Georgopoulos, Martin Rünz, Lourdes Agapito, Matthias Nießner*
39. Zero-Shot Text-to-Parameter Translation for Game Character Auto-Creation, *Rui Zhao, Wei Li, Zhipeng Hu, Lincheng Li, Zhengxia Zou, Zhenwei Shi, Changjie Fan*
40. Learning Locally Editable Virtual Humans, *Hsuan-I Ho, Lixin Xue, Jie Song, Otmar Hilliges*
41. Auto-CARD: Efficient and Robust Codec Avatar Driving for Real-Time Mobile Telepresence, *Yonggan Fu, Yuecheng Li, Chenghui Li, Jason Saragih, Peizhao Zhang, Xiaoliang Dai, Yingyan (Celine) Lin*
42. Ham2Pose: Animating Sign Language Notation Into Pose Sequences, *Rotem Shalev Arkushin, Amit Moryossef, Ohad Fried*
43. PointAvatar: Deformable Point-Based Head Avatars From Videos, *Yufeng Zheng, Wang Yifan, Gordon Wetzstein, Michael J. Black, Otmar Hilliges*
44. PAniC-3D: Stylized Single-View 3D Reconstruction From Portraits of Anime Characters, *Shuhong Chen, Kevin Zhang, Yichun Shi, Heng Wang, Yiheng Zhu, Guoxian Song, Sizhe An, Janus Kristjansson, Xiao Yang, Matthias Zwicker*
45. HandNeRF: Neural Radiance Fields for Animatable Interacting Hands, *Zhiyang Guo, Wengang Zhou, Min Wang, Li Li, Houqiang Li*
46. VGFlow: Visibility Guided Flow Network for Human Reposing, *Rishabh Jain, Krishna Kumar Singh, Mayur Hemani, Jingwan Lu, Mausoom Sarkar, Duygu Ceylan, Balaji Krishnamurthy*
47. Clothed Human Performance Capture With a Double-Layer Neural Radiance Fields, *Kangkan Wang, Guofeng Zhang, Suxu Cong, Jian Yang*
48. POEM: Reconstructing Hand in a Point Embedded Multi-View Stereo, *Lixin Yang, Jian Xu, Licheng Zhong, Xinyu Zhan, Zhicheng Wang, Kejian Wu, Cewu Lu*
49. FlexNeRF: Photorealistic Free-Viewpoint Rendering of Moving Humans From Sparse Views, *Vinoj Jayasundara, Amit Agrawal, Nicolas Heron, Abhinav Shrivastava, Larry S. Davis*
50. Flow Supervision for Deformable NeRF, *Chaoyang Wang, ✨ Lachlan Ewen MacDonald, László A. Jeni, Simon Lucey*
51. Building Rearticulable Models for Arbitrary 3D Objects From 4D Point Clouds, *Shaowei Liu, Saurabh Gupta, Shenlong Wang*
52. Implicit 3D Human Mesh Recovery Using Consistency With Pose and Shape From Unseen-View, *Hanbyel Cho, Yooshin Cho, Jaesung Ahn, Junmo Kim*
53. One-Stage 3D Whole-Body Mesh Recovery With Component Aware Transformer, *Jing Lin, Ailing Zeng, Haoqian Wang, Lei Zhang, Yu Li*
54. Im2Hands: Learning Attentive Implicit Representation of Interacting Two-Hand Shapes, *Jihyun Lee, Minhyuk Sung, Honggyu Choi, Tae-Kyun Kim*
55. FLEX: Full-Body Grasping Without Full-Body Grasps, *Purva Tendulkar, Dídac Surís, Carl Vondrick*
56. DexArt: Benchmarking Generalizable Dexterous Manipulation With Articulated Objects, *Chen Bao, Helin Xu, Yuzhe Qin, Xiaolong Wang*
57. CARTO: Category and Joint Agnostic Reconstruction of ARTiculated Objects, *Nick Heppert, Muhammad Zubair Irshad, Sergey Zakharov, Katherine Liu, Rares Andrei Ambrus, Jeannette Bohg, Abhinav Valada, Thomas Kollar*
58. CIRCLE: Capture in Rich Contextual Environments, *João Pedro Araújo, Jiaman Li, Karthik Vetrivel, Rishi Agarwal, Jiajun Wu, Deepak Gopinath, Alexander William Clegg, Karen Liu*
59. Decoupling Human and Camera Motion From Videos in the Wild, *Vickie Ye, Georgios Pavlakos, Jitendra Malik, Angjoo Kanazawa*
60. GarmentTracking: Category-Level Garment Pose Tracking, *Han Xue, Wenqiang Xu, Jieyi Zhang, Tutian Tang, Yutong Li, Wenxin Du, Ruolin Ye, Cewu Lu*
61. Hierarchical Temporal Transformer for 3D Hand Pose Estimation and Action Recognition From Egocentric RGB Videos, *Yilin Wen, Hao Pan, Lei Yang, Jia Pan, Taku Komura, Wenping Wang*
62. PSVT: End-to-End Multi-Person 3D Pose and Shape Estimation With Progressive Video Transformers, *Zhongwei Qiu, Qiansheng Yang, Jian Wang, Haocheng Feng, Junyu Han, Errui Ding, Chang Xu, Dongmei Fu, Jingdong Wang*

63. Delving Into Discrete Normalizing Flows on $SO(3)$ Manifold for Probabilistic Rotation Modeling, *Yulin Liu, Haoran Liu, Yingda Yin, Yang Wang, Baoquan Chen, He Wang*
64. 3D-POP – An Automated Annotation Approach to Facilitate Markerless 2D-3D Tracking of Freely Moving Birds With Marker-Based Motion Capture, *Hemal Naik, Alex Hoi Hang Chan, Junran Yang, Mathilde Delacoux, Iain D. Couzin, Fumihiko Kano, Máté Nagy*
65. TTA-COPE: Test-Time Adaptation for Category-Level Object Pose Estimation, *Taeyeop Lee, Jonathan Tremblay, Valts Blukis, Bowen Wen, Byeong-Uk Lee, Inkyu Shin, Stan Birchfield, In So Kweon, Kuk-Jin Yoon*
66. Markerless Camera-to-Robot Pose Estimation via Self-Supervised Sim-to-Real Transfer, *Jingpei Lu, Florian Richter, Michael C. Yip*
67. SMOC-Net: Leveraging Camera Pose for Self-Supervised ✧ Monocular Object Pose Estimation, *Tao Tan, Qiulei Dong*
68. IMP: Iterative Matching and Pose Estimation With Adaptive Pooling, *Fei Xue, Ignas Budvytis, Roberto Cipolla*
69. Self-Supervised Representation Learning for CAD, *Benjamin T. Jones, Michael Hu, Milin Kodnongbua, Vladimir G. Kim, Adriana Schulz*
70. Few-Shot Geometry-Aware Keypoint Localization, *Xingzhe He, Gaurav Bharaj, David Ferman, Helge Rhodin, Pablo Garrido*
71. SparsePose: Sparse-View Camera Pose Regression and Refinement, *Samarth Sinha, Jason Y. Zhang, Andrea Tagliasacchi, Igor Gilitschenski, David B. Lindell*
72. A Large-Scale Homography Benchmark, *Daniel Barath, Dmytro Mishkin, Michal Polic, Wolfgang Förstner, Jiri Matas*
73. Learning Geometric-Aware Properties in 2D Representation Using Lightweight CAD Models, or Zero Real 3D Pairs, *Pattaramanee Arsomngern, Sarana Nutanong, Supasorn Suwajanakorn*
74. AutoRecon: Automated 3D Object Discovery and ✧ Reconstruction, *Yuang Wang, Xingyi He, Sida Peng, Haotong Lin, Hujun Bao, Xiaowei Zhou*
75. Multi-Sensor Large-Scale Dataset for Multi-View 3D Reconstruction, *Oleg Voynov, Gleb Bobrovskikh, Pavel Karpyshev, Saveliy Galochkin, Andrei-Timotei Ardelean, Arseniy Bozhenko, Ekaterina Karmanova, Pavel Kopanev, Yaroslav Labutin-Rymsho, Ruslan Rakhimov, Aleksandr Safin, Valerii Serpiva, Alexey Artemov, Evgeny Burnaev, Dzmitry Tsetserukou, Denis Zorin*
76. NeurOCS: Neural NOCS Supervision for Monocular 3D Object Localization, *Zhixiang Min, Bingbing Zhuang, Samuel Schulter, Buyu Liu, Enrique Dunn, Manmohan Chandraker*
77. Self-Supervised Super-Plane for Neural 3D Reconstruction, *Botao Ye, Sifei Liu, Xueting Li, Ming-Hsuan Yang*
78. PlaneDepth: Self-Supervised Depth Estimation via Orthogonal Planes, *Ruoyu Wang, Zehao Yu, Shenghua Gao*
79. Single View Scene Scale Estimation Using Scale Field, *Byeong-Uk Lee, Jianming Zhang, Yannick Hold-Geoffroy, In So Kweon*
80. 3D Line Mapping Revisited, *Shaohui Liu, Yifan Yu, Rémi Pautrat, ✧ Marc Pollefeys, Viktor Larsson*
81. Inverting the Imaging Process by Learning an Implicit Camera Model, *Xin Huang, Qi Zhang, Ying Feng, Hongdong Li, Qing Wang*
82. SfM-TTR: Using Structure From Motion for Test-Time Refinement of Single-View Depth Networks, *Sergio Izquierdo, Javier Civera*
83. iDisc: Internal Discretization for Monocular Depth Estimation, *Luigi Piccinelli, Christos Sakaridis, Fisher Yu*
84. DC²: Dual-Camera Defocus Control by Learning to Refocus, *Hadi Alzayer, Abdullah Abuolaim, Leung Chun Chan, Yang Yang, Ying Chen Lou, Jia-Bin Huang, Abhishek Kar*
85. A Practical Stereo Depth System for Smart Glasses, *Jialiang Wang, Daniel Scharstein, Akash Bapat, Kevin Blackburn-Matzen, Matthew Yu, Jonathan Lehman, Suhil Alsian, Yanghan Wang, Sam Tsai, Jan-Michael Frahm, Zijian He, Peter Vajda, Michael F. Cohen, Matt Uyttendaele*
86. GeoMVSNet: Learning Multi-View Stereo With Geometry Perception, *Zhe Zhang, Rui Peng, Yuxi Hu, Ronggang Wang*
87. DINN360: Deformable Invertible Neural Network for Latitude-Aware 360° Image Rescaling, *Yichen Guo, Mai Xu, Lai Jiang, Leonid Sigal, Yunjin Chen*
88. OmniVidar: Omnidirectional Depth Estimation From Multi-Fisheye Images, *Sheng Xie, Daochuan Wang, Yun-Hui Liu*
89. Learning to Fuse Monocular and Multi-View Cues for Multi-Frame Depth Estimation in Dynamic Scenes, *Rui Li, Dong Gong, Wei Yin, Hao Chen, Yu Zhu, Kaixuan Wang, Xiaozhi Chen, Jinqiu Sun, Yanning Zhang*
90. Modality-Invariant Visual Odometry for Embodied Vision, *Marius Memmel, Roman Bachmann, Amir Zamir*
91. VL-SAT: Visual-Linguistic Semantics Assisted Training for 3D ✧ Semantic Scene Graph Prediction in Point Cloud, *Ziqin Wang, Bowen Cheng, Lichen Zhao, Dong Xu, Yang Tang, Lu Sheng*
92. CAPE: Camera View Position Embedding for Multi-View 3D Object Detection, *Kaixin Xiong, Shi Gong, Xiaoqing Ye, Xiao Tan, Ji Wan, Errui Ding, Jingdong Wang, Xiang Bai*
93. AeDet: Azimuth-Invariant Multi-View 3D Object Detection, *Chengjian Feng, Zequn Jie, Yujie Zhong, Xiangxiang Chu, Lin Ma*
94. Object Detection With Self-Supervised Scene Adaptation, *Zekun Zhang, Minh Hoai*
95. Understanding the Robustness of 3D Object Detection With Bird's-Eye-View Representations in Autonomous Driving, *Zijian Zhu, Yichi Zhang, Hai Chen, Yinpeng Dong, Shu Zhao, Wenbo Ding, Jiachen Zhong, Shibao Zheng*
96. BEVHeight: A Robust Framework for Vision-Based Roadside 3D Object Detection, *Lei Yang, Kaicheng Yu, Tao Tang, Jun Li, Kun Yuan, Li Wang, Xinyu Zhang, Peng Chen*
97. Uncertainty-Aware Vision-Based Metric Cross-View Geolocalization, *Florian Fervers, Sebastian Bullinger, Christoph Bodensteiner, Michael Arens, Rainer Stiefelhagen*
98. OrienterNet: Visual Localization in 2D Public Maps With Neural Matching, *Paul-Edouard Sarlin, Daniel DeTone, Tsun-Yi Yang, Armen Avetisyan, Julian Straub, Tomasz Malisiewicz, Samuel Rota Bulò, Richard Newcombe, Peter Kotschieder, Vasileios Balntas*
99. MSMDFusion: Fusing LiDAR and Camera at Multiple Scales With Multi-Depth Seeds for 3D Object Detection, *Yang Jiao, Zequn Jie, Shaoxiang Chen, Jingjing Chen, Lin Ma, Yu-Gang Jiang*
100. Virtual Sparse Convolution for Multimodal 3D Object Detection, *Hai Wu, Chenglu Wen, Shaoshuai Shi, Xin Li, Cheng Wang*
101. Optimal Transport Minimization: Crowd Localization on Density ✧ Maps for Semi-Supervised Counting, *Wei Lin, Antoni B. Chan*

102. VoxelNeXt: Fully Sparse VoxelNet for 3D Object Detection and Tracking, *Yukang Chen, Jianhui Liu, Xiangyu Zhang, Xiaojuan Qi, Jiaya Jia*
103. GraVoS: Voxel Selection for 3D Point-Cloud Detection, *Oren Shrouf, Yizhak Ben-Shabat, Ayellet Tal*
104. MSeg3D: Multi-Modal 3D Semantic Segmentation for Autonomous Driving, *Jiale Li, Hang Dai, Hao Han, Yong Ding*
105. LaserMix for Semi-Supervised LiDAR Semantic Segmentation, ✧ *Lingdong Kong, Jiawei Ren, Liang Pan, Ziwei Liu*
106. Implicit Surface Contrastive Clustering for LiDAR Point Clouds, *Zaiwei Zhang, Min Bai, Erran Li*
107. Semi-Weakly Supervised Object Kinematic Motion Prediction, *Gengxin Liu, Qian Sun, Haibin Huang, Chongyang Ma, Yulan Guo, Li Yi, Hui Huang, Ruizhen Hu*
108. PartSLIP: Low-Shot Part Segmentation for 3D Point Clouds via Pretrained Image-Language Models, *Minghua Liu, Yinhao Zhu, Hong Cai, Shizhong Han, Zhan Ling, Fatih Porikli, Hao Su*
109. Learning Weather-General and Weather-Specific Features for Image Restoration Under Multiple Adverse Weather Conditions, *Yurui Zhu, Tianyu Wang, Xueyang Fu, Xuanyu Yang, Xin Guo, Jifeng Dai, Yu Qiao, Xiaowei Hu*
110. Geometry and Uncertainty-Aware 3D Point Cloud Class-Incremental Semantic Segmentation, *Yuwei Yang, Munawar Hayat, Zhao Jin, Chao Ren, Yinjie Lei*
111. Learning 3D Representations From 2D Pre-Trained Models via Image-to-Point Masked Autoencoders, *Renrui Zhang, Lihui Wang, Yu Qiao, Peng Gao, Hongsheng Li*
112. ToThePoint: Efficient Contrastive Learning of 3D Point Clouds via Recycling, *Xinglin Li, Jiajing Chen, Jinhui Ouyang, Hanhui Deng, Senem Velipasalar, Di Wu*
113. PointDistiller: Structured Knowledge Distillation Towards Efficient and Compact 3D Detection, *Linfeng Zhang, Runpei Dong, Hung-Shuo Tai, Kaisheng Ma*
114. PointConvFormer: Revenge of the Point-Based Convolution, *Wenxuan Wu, Li Fuxin, Qi Shan*
115. Self-Positioning Point-Based Transformer for Point Cloud Understanding, *Jinyoung Park, Sanghyeok Lee, Sihyeon Kim, Yonyang Xiong, Hyunwoo J. Kim*
116. PointClustering: Unsupervised Point Cloud Pre-Training Using ✧ *Transformation Invariance in Clustering, Fuchen Long, Ting Yao, Zhaofan Qiu, Lusong Li, Tao Mei*
117. Neural Intrinsic Embedding for Non-Rigid Point Cloud Matching, *Puhua Jiang, Mingze Sun, Ruqi Huang*
118. HGNet: Learning Hierarchical Geometry From Points, Edges, and Surfaces, *Ting Yao, Yehao Li, Yingwei Pan, Tao Mei*
119. LP-DIF: Learning Local Pattern-Specific Deep Implicit Function for 3D Objects and Scenes, *Meng Wang, Yu-Shen Liu, Yue Gao, Kanle Shi, Yi Fang, Zhizhong Han*
120. Conjugate Product Graphs for Globally Optimal 2D-3D Shape Matching, *Paul Roetzer, Zorah Löhner, Florian Bernard*
121. UTM: A Unified Multiple Object Tracking Model With Identity-Aware Feature Enhancement, *Sisi You, Hantao Yao, Bing-Kun Bao, Changsheng Xu*
122. Learning Rotation-Equivariant Features for Visual Correspondence, *Jongmin Lee, Byungjin Kim, Seungwook Kim, Minsu Cho*
123. Adaptive Spot-Guided Transformer for Consistent Local Feature Matching, *Jiahuan Yu, Jiahao Chang, Jianfeng He, Tianzhu Zhang, Jiyang Yu, Feng Wu*
124. PMatch: Paired Masked Image Modeling for Dense Geometric Matching, *Shengjie Zhu, Xiaoming Liu*
125. Iterative Geometry Encoding Volume for Stereo Matching, *Gangwei Xu, Xianqi Wang, Xiaohuan Ding, Xin Yang*
126. Adaptive Annealing for Robust Geometric Estimation, *Chitturi Sidhartha, Lalit Manam, Venu Madhav Govindu*
127. Tangentially Elongated Gaussian Belief Propagation for Event-Based Incremental Optical Flow Estimation, *Jun Nagata, Yusuke Sekikawa*
128. Robust and Scalable Gaussian Process Regression and Its Applications, *Yifan Lu, Jiayi Ma, Leyuan Fang, Xin Tian, Junjun Jiang*
129. BEV-Guided Multi-Modality Fusion for Driving Perception, *Yunze Man, Liang-Yan Gui, Yu-Xiong Wang*
130. HumanBench: Towards General Human-Centric Perception With Projector Assisted Pretraining, *Shixiang Tang, Cheng Chen, Qingsong Xie, Meilin Chen, Yizhou Wang, Yuanzheng Ci, Lei Bai, Feng Zhu, Haiyang Yang, Li Yi, Rui Zhao, Wanli Ouyang*
131. Think Twice Before Driving: Towards Scalable Decoders for End-to-End Autonomous Driving, *Xiaosong Jia, Penghao Wu, Li Chen, Jiangwei Xie, Conghui He, Junchi Yan, Hongyang Li*
132. ProphNet: Efficient Agent-Centric Motion Forecasting With ✧ *Anchor-Informed Proposals, Xishun Wang, Tong Su, Fang Da, Xiaodong Yang*
133. StarCraftImage: A Dataset for Prototyping Spatial Reasoning Methods for Multi-Agent Environments, *Sean Kulinski, Nicholas R. Waytowich, James Z. Hare, David I. Inouye*
134. Stimulus Verification Is a Universal and Effective Sampler in Multi-Modal Human Trajectory Prediction, *Jianhua Sun, Yuxuan Li, Liang Chai, Cewu Lu*
135. PyPose: A Library for Robot Learning With Physics-Based Optimization, *Chen Wang, Dasong Gao, Kuan Xu, Junyi Geng, Yaoyu Hu, Yuheng Qiu, Bowen Li, Fan Yang, Brady Moon, Abhinav Pandey, Aryan, Jiahe Xu, Tianhao Wu, Haonan He, Daning Huang, Zhongqiang Ren, Shibo Zhao, Taimeng Fu, Pranay Reddy, Xiao Lin, Wenshan Wang, Jingnan Shi, Rajat Talak, Kun Cao, Yi Du, Han Wang, Huai Yu, Shanzhao Wang, Siyu Chen, Ananth Kashyap, Rohan Bandaru, Karthik Dantu, Jiajun Wu, Lihua Xie, Luca Carlone, Marco Hutter, Sebastian Scherer*
136. Source-Free Adaptive Gaze Estimation by Uncertainty Reduction, *Xin Cai, Jiabei Zeng, Shiguang Shan, Xilin Chen*
137. Camouflaged Object Detection With Feature Decomposition and Edge Reconstruction, *Chunming He, Kai Li, Yachao Zhang, Longxiang Tang, Yulun Zhang, Zhenhua Guo, Xiu Li*
138. MOTRv2: Bootstrapping End-to-End Multi-Object Tracking by Pretrained Object Detectors, *Yuang Zhang, Tiancai Wang, Xiangyu Zhang*
139. Dynamic Aggregated Network for Gait Recognition, *Kang Ma, Ying Fu, Dezhi Zheng, Chunshui Cao, Xuecai Hu, Yongzhen Huang*
140. Feature Representation Learning With Adaptive Displacement Generation and Transformer Fusion for Micro-Expression Recognition, *Zhijun Zhai, Jianhui Zhao, Chengjiang Long, Wenju Xu, Shuangjiang He, Huijuan Zhao*

141. MetaPortrait: Identity-Preserving Talking Head Generation With Fast Personalized Adaptation, *Bowen Zhang, Chenyang Qi, Pan Zhang, Bo Zhang, HsiangTao Wu, Dong Chen, Qifeng Chen, Yong Wang, Fang Wen*
142. FLAG3D: A 3D Fitness Activity Dataset With Language Instruction, *Yansong Tang, Jinpeng Liu, Aoyang Liu, Bin Yang, Wenxun Dai, Yongming Rao, Jiwen Lu, Jie Zhou, Xiu Li*
143. TranSG: Transformer-Based Skeleton Graph Prototype Contrastive Learning With Structure-Trajectory Prompted Reconstruction for Person Re-Identification, *Haocong Rao, Chunyan Miao*
144. NeMo: Learning 3D Neural Motion Fields From Multiple Video ✧ Instances of the Same Action, *Kuan-Chieh Wang, Zhenzhen Weng, Maria Xenochristou, João Pedro Araújo, Jeffrey Gu, Karen Liu, Serena Yeung*
145. Unsupervised Space-Time Network for Temporally-Consistent ✧ Segmentation of Multiple Motions, *Etienne Meunier, Patrick Bouthemy*
146. Deep Polarization Reconstruction With PDAVIS Events, *Haiyang Mei, Zuowen Wang, Xin Yang, Xiaopeng Wei, Tobi Delbruck*
147. Range-Nullspace Video Frame Interpolation With Focalized Motion Estimation, *Zhiyang Yu, Yu Zhang, Dongqing Zou, Xijun Chen, Jimmy S. Ren, Shunqing Ren*
148. Exploring Motion Ambiguity and Alignment for High-Quality Video Frame Interpolation, *Kun Zhou, Wenbo Li, Xiaoguang Han, Jiangbo Lu*
149. 1000 FPS HDR Video With a Spike-RGB Hybrid Camera, *Yakun Chang, Chu Zhou, Yuchen Hong, Liwen Hu, Chao Xu, Tiejun Huang, Boxin Shi*
150. Deep Discriminative Spatial and Temporal Network for Efficient Video Deblurring, *Jinshan Pan, Boming Xu, Jiangxin Dong, Jianjun Ge, Jinhui Tang*
151. Gated Multi-Resolution Transfer Network for Burst Restoration and Enhancement, *Nancy Mehta, Akshay Dudhane, Subrahmanyam Murala, Syed Waqas Zamir, Salman Khan, Fahad Shahbaz Khan*
152. A Unified HDR Imaging Method With Pixel and Patch Level, *Qingsen Yan, Weiye Chen, Song Zhang, Yu Zhu, Jinqiu Sun, Yanning Zhang*
153. BiasBed – Rigorous Texture Bias Evaluation, *Nikolai Kalischek, Rodrigo Caye Daudt, Torben Peters, Reinhard Furrer, Jan D. Wegner, Konrad Schindler*
154. Learning a Practical SDR-to-HDRTV Up-Conversion Using New Dataset and Degradation Models, *Cheng Guo, Leidong Fan, Ziyu Xue, Xiuhua Jiang*
155. Learning a Deep Color Difference Metric for Photographic Images, *Haoyu Chen, Zhihua Wang, Yang Yang, Qilin Sun, Kede Ma*
156. Learning a Simple Low-Light Image Enhancer From Paired Low-Light Instances, *Zhenqi Fu, Yan Yang, Xiaotong Tu, Yue Huang, Xinghao Ding, Kai-Kuang Ma*
157. Residual Degradation Learning Unfolding Framework With Mixing Priors Across Spectral and Spatial for Compressive Spectral Imaging, *Yubo Dong, Dahua Gao, Tian Qiu, Yuyan Li, Minxi Yang, Guangming Shi*
158. Toward Stable, Interpretable, and Lightweight Hyperspectral Super-Resolution, *Wen-jin Guo, Weiying Xie, Kai Jiang, Yunsong Li, Jie Lei, Leyuan Fang*
159. RIDCP: Revitalizing Real Image Dehazing via High-Quality Codebook Priors, *Rui-Qi Wu, Zheng-Peng Duan, Chun-Le Guo, Zhi Chai, Chongyi Li*
160. Robust Unsupervised StyleGAN Image Restoration, *Yohan Poirier-Ginter, Jean-François Lalonde*
161. Quality-Aware Pre-Trained Models for Blind Image Quality Assessment, *Kai Zhao, Kun Yuan, Ming Sun, Mading Li, Xing Wen*
162. Learning to Exploit the Sequence-Specific Prior Knowledge for Image Processing Pipelines Optimization, *Haina Qin, Longfei Han, Weihua Xiong, Juan Wang, Wentao Ma, Bing Li, Weiming Hu*
163. Multi-Realism Image Compression With a Conditional Generator, *Eirikur Agustsson, David Minnen, George Toderici, Fabian Mentzer*
164. RGB No More: Minimally-Decoded JPEG Vision Transformers, *Jeongsoo Park, Justin Johnson*
165. Kernel Aware Resampler, *Michael Bernasconi, Abdelaziz Djelouah, Farnood Salehi, Markus Gross, Christopher Schroers*
166. Spatial-Frequency Mutual Learning for Face Super-Resolution, *Chenyang Wang, Junjun Jiang, Zhiwei Zhong, Xianming Liu*
167. Activating More Pixels in Image Super-Resolution Transformer, *Xiangyu Chen, Xintao Wang, Jiantao Zhou, Yu Qiao, Chao Dong*
168. Omni Aggregation Networks for Lightweight Image Super-Resolution, *Hang Wang, Xuanhong Chen, Bingbing Ni, Yutian Liu, Jinfan Liu*
169. Towards Artistic Image Aesthetics Assessment: A Large-Scale Dataset and a New Method, *Ran Yi, Haoyuan Tian, Zhihao Gu, Yu-Kun Lai, Paul L. Rosin*
170. RWSC-Fusion: Region-Wise Style-Controlled Fusion Network for the Prohibited X-Ray Security Image Synthesis, *Luwen Duan, Min Wu, Lijian Mao, Jun Yin, Jianping Xiong, Xi Li*
171. Efficient Scale-Invariant Generator With Column-Row Entangled Pixel Synthesis, *Thuan Hoang Nguyen, Thanh Van Le, Anh Tran*
172. Masked and Adaptive Transformer for Exemplar Based Image Translation, *Chang Jiang, Fei Gao, Biao Ma, Yuhao Lin, Nannan Wang, Gang Xu*
173. SmartBrush: Text and Shape Guided Object Inpainting With ✧ Diffusion Model, *Shaoan Xie, Zhifei Zhang, Zhe Lin, Tobias Hinz, Kun Zhang*
174. Neural Transformation Fields for Arbitrary-Styled Font Generation, *Bin Fu, Junjun He, Jianjun Wang, Yu Qiao*
175. Referring Image Matting, *Jizhizi Li, Jing Zhang, Dacheng Tao*
176. Handwritten Text Generation From Visual Archetypes, *Vittorio Pippi, Silvia Cascianelli, Rita Cucchiara*
177. SceneComposer: Any-Level Semantic Image Synthesis, *Yu ✧ Zeng, Zhe Lin, Jianming Zhang, Qing Liu, John Collomosse, Jason Kuen, Vishal M. Patel*
178. Affordance Diffusion: Synthesizing Hand-Object Interactions, *Yufei Ye, Xueting Li, Abhinav Gupta, Shalini De Mello, Stan Birchfield, Jiaming Song, Shubham Tulsiani, Sifei Liu*
179. LayoutDiffusion: Controllable Diffusion Model for Layout-to-Image Generation, *Guangcong Zheng, Xianpan Zhou, Xuwei Li, Zhongang Qi, Ying Shan, Xi Li*
180. DreamBooth: Fine Tuning Text-to-Image Diffusion Models for ✧ Subject-Driven Generation, *Nataniel Ruiz, Yanzhen Li, Varun Jampani, Yael Pritch, Michael Rubinstein, Kfir Aberman*

181. GLIGEN: Open-Set Grounded Text-to-Image Generation, *Yuheng Li, Haotian Liu, Qingyang Wu, Fangzhou Mu, Jianwei Yang, Jianfeng Gao, Chunyuan Li, Yong Jae Lee*
182. Safe Latent Diffusion: Mitigating Inappropriate Degeneration in Diffusion Models, *Patrick Schramowski, Manuel Brack, Björn Deiseroth, Kristian Kersting*
183. EDICT: Exact Diffusion Inversion via Coupled Transformations, *Bram Wallace, Akash Gokul, Nikhil Naik*
184. Solving 3D Inverse Problems Using Pre-Trained 2D Diffusion Models, *Hyungjin Chung, Dohoon Ryu, Michael T. McCann, Marc L. Klasky, Jong Chul Ye*
185. Diffusion Probabilistic Model Made Slim, *Xingyi Yang, Daquan Zhou, Jiashi Feng, Xinchao Wang*
186. Align Your Latents: High-Resolution Video Synthesis With Latent Diffusion Models, *Andreas Blattmann, Robin Rombach, Huan Ling, Tim Dockhorn, Seung Wook Kim, Sanja Fidler, Karsten Kreis*
187. Binary Latent Diffusion, *Ze Wang, Jiang Wang, Zicheng Liu, Qiang Qiu*
188. Semi-Supervised Video Inpainting With Cycle Consistency Constraints, *Zhiliang Wu, Hanyu Xuan, Changchang Sun, Weili Guan, Kang Zhang, Yan Yan*
189. Towards Accurate Image Coding: Improved Autoregressive ✧ Image Generation With Dynamic Vector Quantization, *Mengqi Huang, Zhendong Mao, Zhuowei Chen, Yongdong Zhang*
190. Large-Capacity and Flexible Video Steganography via Invertible Neural Network, *Chong Mou, Youmin Xu, Jiechong Song, Chen Zhao, Bernard Ghanem, Jian Zhang*
191. Neural Video Compression With Diverse Contexts, *Jiahao Li, Bin Li, Yan Lu*
192. Efficient Semantic Segmentation by Altering Resolutions for Compressed Videos, *Yubin Hu, Yuze He, Yanghao Li, Jisheng Li, Yuxing Han, Jiangtao Wen, Yong-Jin Liu*
193. Structured Sparsity Learning for Efficient Video Super-Resolution, *Bin Xia, Jingwen He, Yulun Zhang, Yitong Wang, Yapeng Tian, Wenming Yang, Luc Van Gool*
194. DisCo-CLIP: A Distributed Contrastive Loss for Memory Efficient ✧ CLIP Training, *Yihao Chen, Xianbiao Qi, Jianan Wang, Lei Zhang*
195. Boost Vision Transformer With GPU-Friendly Sparsity and Quantization, *Chong Yu, Tao Chen, Zhongxue Gan, Jiayuan Fan*
196. All Are Worth Words: A ViT Backbone for Diffusion Models, *Fan Bao, Shen Nie, Kaiwen Xue, Yue Cao, Chongxuan Li, Hang Su, Jun Zhu*
197. Sparsifiner: Learning Sparse Instance-Dependent Attention for Efficient Vision Transformers, *Cong Wei, Brendan Duke, Ruowei Jiang, Parham Aarabi, Graham W. Taylor, Florian Shkurti*
198. DropKey for Vision Transformer, *Bonan Li, Yinhan Hu, Xuecheng Nie, Congying Han, Xiangjian Jiang, Tiande Guo, Luoqi Liu*
199. Seeing Beyond the Brain: Conditional Diffusion Model With Sparse Masked Modeling for Vision Decoding, *Zijiao Chen, Jiaxin Qing, Tiange Xiang, Wan Lin Yue, Juan Helen Zhou*
200. ResFormer: Scaling ViTs With Multi-Resolution Training, *Rui Tian, Zuxuan Wu, Qi Dai, Han Hu, Yu Qiao, Yu-Gang Jiang*
201. Stare at What You See: Masked Image Modeling Without Reconstruction, *Hongwei Xue, Peng Gao, Hongyang Li, Yu Qiao, Hao Sun, Houqiang Li, Jiebo Luo*
202. Mixed Autoencoder for Self-Supervised Visual Representation Learning, *Kai Chen, Zhili Liu, Lanqing Hong, Hang Xu, Zhenguo Li, Dit-Yan Yeung*
203. Shape-Erased Feature Learning for Visible-Infrared Person Re-Identification, *Jiawei Feng, Ancong Wu, Wei-Shi Zheng*
204. G-MSM: Unsupervised Multi-Shape Matching With Graph-Based Affinity Priors, *Marvin Eisenberger, Aysim Toker, Laura Leal-Taixé, Daniel Cremers*
205. Efficient Mask Correction for Click-Based Interactive Image Segmentation, *Fei Du, Jianlong Yuan, Zhibin Wang, Fan Wang*
206. Prototype-Based Embedding Network for Scene Graph Generation, *Chaofan Zheng, Xinyu Lyu, Lianli Gao, Bo Dai, Jingkuan Song*
207. Graph Representation for Order-Aware Visual Transformation, *Yue Qiu, Yanjun Sun, Fumiya Matsuzawa, Kenji Iwata, Hirokatsu Kataoka*
208. Unbiased Scene Graph Generation in Videos, *Sayak Nag, Kyle Min, Subarna Tripathi, Amit K. Roy-Chowdhury*
209. Recurrence Without Recurrence: Stable Video Landmark Detection With Deep Equilibrium Models, *Paul Micaelli, Arash Vahdat, Hongxu Yin, Jan Kautz, Pavlo Molchanov*
210. VideoTrack: Learning to Track Objects via Video Transformer, *Fei Xie, Lei Chu, Jiahao Li, Yan Lu, Chao Ma*
211. Breaking the "Object" in Video Object Segmentation, *Pavel Tokmakov, Jie Li, Adrien Gaidon*
212. Hierarchical Semantic Contrast for Scene-Aware Video Anomaly Detection, *Shengyang Sun, Xiaojin Gong*
213. Mask-Free Video Instance Segmentation, *Lei Ke, Martin Danelljan, Henghui Ding, Yu-Wing Tai, Chi-Keung Tang, Fisher Yu*
214. Hierarchical Neural Memory Network for Low Latency Event Processing, *Ryuhei Hamaguchi, Yasutaka Furukawa, Masaki Onishi, Ken Sakurada*
215. Unifying Short and Long-Term Tracking With Graph Hierarchies, *Orcun Cetintas, Guillem Brasó, Laura Leal-Taixé*
216. Towards End-to-End Generative Modeling of Long Videos With Memory-Efficient Bidirectional Transformers, *Jaehoon Yoo, Semin Kim, Doyup Lee, Chiheon Kim, Seunghoon Hong*
217. An Empirical Study of End-to-End Video-Language Transformers With Masked Visual Modeling, *Tsu-Jui Fu, Linjie Li, Zhe Gan, Kevin Lin, William Yang Wang, Lijuan Wang, Zicheng Liu*
218. Egocentric Audio-Visual Object Localization, *Chao Huang, Yapeng Tian, Anurag Kumar, Chenliang Xu*
219. AVFormer: Injecting Vision Into Frozen Speech Models for Zero-Shot AV-ASR, *Paul Hongsuck Seo, Arsha Nagrani, Cordelia Schmid*
220. A Light Weight Model for Active Speaker Detection, *Junhua Liao, Haihan Duan, Kanghui Feng, Wanbing Zhao, Yanbing Yang, Liangyin Chen*
221. Dense-Localizing Audio-Visual Events in Untrimmed Videos: A Large-Scale Benchmark and Baseline, *Tiantian Geng, Teng Wang, Jinming Duan, Runmin Cong, Feng Zheng*
222. Video Test-Time Adaptation for Action Recognition, *Wei Lin, Muhammad Jehanzeb Mirza, Mateusz Kozinski, Horst Possegger, Hilde Kuehne, Horst Bischof*
223. Unified Keypoint-Based Action Recognition Framework via Structured Keypoint Pooling, *Ryo Hachiuma, Fumiaki Sato, Taiki Sekii*

224. Object Discovery From Motion-Guided Tokens, *Zhipeng Bao, Pavel Tokmakov, Yu-Xiong Wang, Adrien Gaidon, Martial Hebert*
225. Open Set Action Recognition via Multi-Label Evidential Learning, *Chen Zhao, Dawei Du, Anthony Hoogs, Christopher Funk*
226. PivoTAL: Prior-Driven Supervision for Weakly-Supervised Temporal Action Localization, *Mamshad Nayeem Rizve, Gaurav Mittal, Ye Yu, Matthew Hall, Sandra Sajeev, Mubarak Shah, Mei Chen*
227. Improving Weakly Supervised Temporal Action Localization by Bridging Train-Test Gap in Pseudo Labels, *Jingqiu Zhou, Linjiang Huang, Liang Wang, Si Liu, Hongsheng Li*
228. Are Binary Annotations Sufficient? Video Moment Retrieval via Hierarchical Uncertainty-Based Active Learning, *Wei Ji, Renjie Liang, Zhedong Zheng, Wenqiao Zhang, Shengyu Zhang, Juncheng Li, Mengze Li, Tat-seng Chua*
229. Query-Dependent Video Representation for Moment Retrieval and Highlight Detection, *WonJun Moon, Sangeek Hyun, SangUk Park, Dongchan Park, Jae-Pil Heo*
230. Vita-CLIP: Video and Text Adaptive CLIP via Multimodal Prompting, *Syed Talal Wasim, Muzammal Naseer, Salman Khan, Fahad Shahbaz Khan, Mubarak Shah*
231. Towards Generalisable Video Moment Retrieval: Visual-Dynamic Injection to Image-Text Pre-Training, *Dezhao Luo, Jiabo Huang, Shaogang Gong, Hailin Jin, Yang Liu*
232. Hierarchical Video-Moment Retrieval and Step-Captioning, *Abhay Zala, Jaemin Cho, Satwik Kottur, Xilun Chen, Barlas Oguz, Yashar Mehdad, Mohit Bansal*
233. HierVL: Learning Hierarchical Video-Language Embeddings, ✨ *Kumar Ashutosh, Rohit Girdhar, Lorenzo Torresani, Kristen Grauman*
234. Learning Transferable Spatiotemporal Representations From Natural Script Knowledge, *Ziyun Zeng, Yuying Ge, Xihui Liu, Bin Chen, Ping Luo, Shu-Tao Xia, Yixiao Ge*
235. WINNER: Weakly-Supervised hierarchical decomposition and alignment for Spatio-temporal Video grounding, *Mengze Li, Han Wang, Wenqiao Zhang, Jiaxu Miao, Zhou Zhao, Shengyu Zhang, Wei Ji, Fei Wu*
236. Collaborative Static and Dynamic Vision-Language Streams for Spatio-Temporal Video Grounding, *Zihang Lin, Chaolei Tan, Jian-Fang Hu, Zhi Jin, Tiancai Ye, Wei-Shi Zheng*
237. Learning Action Changes by Measuring Verb-Adverb Textual Relationships, *Davide Moltisanti, Frank Keller, Hakan Bilen, Laura Sevilla-Lara*
238. LAVENDER: Unifying Video-Language Understanding As Masked Language Modeling, *Linjie Li, Zhe Gan, Kevin Lin, Chung-Ching Lin, Zicheng Liu, Ce Liu, Lijuan Wang*
239. DeCo: Decomposition and Reconstruction for Compositional Temporal Grounding via Coarse-To-Fine Contrastive Ranking, *Lijin Yang, Quan Kong, Hsuan-Kung Yang, Wadim Kehl, Yoichi Sato, Norimasa Kobori*
240. CVT-SLR: Contrastive Visual-Textual Transformation for Sign Language Recognition With Variational Alignment, *Jiangbin Zheng, Yile Wang, Cheng Tan, Siyuan Li, Ge Wang, Jun Xia, Yidong Chen, Stan Z. Li*
241. Joint Visual Grounding and Tracking With Natural Language Specification, *Li Zhou, Zikun Zhou, Kaige Mao, Zhenyu He*
242. Accelerating Vision-Language Pretraining With Free Language Modeling, *Teng Wang, Yixiao Ge, Feng Zheng, Ran Cheng, Ying Shan, Xiaohu Qie, Ping Luo*
243. CoWs on Pasture: Baselines and Benchmarks for Language-Driven Zero-Shot Object Navigation, *Samir Yitzhak Gadre, Mitchell Wortsman, Gabriel Ilharco, Ludwig Schmidt, Shuran Song*
244. Where We Are and What We're Looking At: Query Based Worldwide Image Geo-Localization Using Hierarchies and Scenes, *Brandon Clark, Alec Kerrigan, Parth Parag Kulkarni, Vicente Vivanco Cepeda, Mubarak Shah*
245. ANetQA: A Large-Scale Benchmark for Fine-Grained Compositional Reasoning Over Untrimmed Videos, *Zhou Yu, Lixiang Zheng, Zhou Zhao, Fei Wu, Jianping Fan, Kui Ren, Jun Yu*
246. MetaCLUE: Towards Comprehensive Visual Metaphors Research, *Arjun R. Akula, Brendan Driscoll, Pradyumna Narayana, Soravit Changpinyo, Zhiwei Jia, Suyash Damle, Garima Pruthi, Sugato Basu, Leonidas Guibas, William T. Freeman, Yuanzhen Li, Varun Jampani*
247. GeoVLN: Learning Geometry-Enhanced Visual Representation With Slot Attention for Vision-and-Language Navigation, *Jingyang Huo, Qiang Sun, Boyan Jiang, Haitao Lin, Yanwei Fu*
248. Being Comes From Not-Being: Open-Vocabulary Text-to-Motion Generation With Wordless Training, *Junfan Lin, Jianlong Chang, Lingbo Liu, Guanbin Li, Liang Lin, Qi Tian, Chang-Wen Chen*
249. LASP: Text-to-Text Optimization for Language-Aware Soft Prompting of Vision & Language Models, *Adrian Bulat, Georgios Tzimiropoulos*
250. Position-Guided Text Prompt for Vision-Language Pre-Training, *Jinpeng Wang, Pan Zhou, Mike Zheng Shou, Shuicheng Yan*
251. Intrinsic Physical Concepts Discovery With Object-Centric Predictive Models, *Qu Tang, Xiangyu Zhu, Zhen Lei, Zhaoxiang Zhang*
252. MAP: Multimodal Uncertainty-Aware Vision-Language Pre-Training Model, *Yatai Ji, Junjie Wang, Yuan Gong, Lin Zhang, Yanru Zhu, Hongfa Wang, Jiaxing Zhang, Tetsuya Sakai, Yujiu Yang*
253. CLAMP: Prompt-Based Contrastive Learning for Connecting Language and Animal Pose, *Xu Zhang, Wen Wang, Zhe Chen, Yufei Xu, Jing Zhang, Dacheng Tao*
254. Teacher-Generated Spatial-Attention Labels Boost Robustness and Accuracy of Contrastive Models, *Yushi Yao, Chang Ye, Junfeng He, Gamaleldin F. Elsayed*
255. DegAE: A New Pretraining Paradigm for Low-Level Vision, ✨ *Yihao Liu, Jingwen He, Jinjin Gu, Xiangtao Kong, Yu Qiao, Chao Dong*
256. RILS: Masked Visual Reconstruction in Language Semantic Space, *Shusheng Yang, Yixiao Ge, Kun Yi, Dian Li, Ying Shan, Xiaohu Qie, Xinggang Wang*
257. Learning Geometry-Aware Representations by Sketching, *Hyundo Lee, Inwoo Hwang, Hyunsung Go, Won-Seok Choi, Kibeom Kim, Byoung-Tak Zhang*
258. SketchXAI: A First Look at Explainability for Human Sketches, *Zhiyu Qu, Yulia Gryaditskaya, Ke Li, Kaiyue Pang, Tao Xiang, Yi-Zhe Song*

259. MAGVLT: Masked Generative Vision-and-Language Transformer, *Sungwoong Kim, Daejin Jo, Donghoon Lee, Jongmin Kim*
260. Zero-Shot Everything Sketch-Based Image Retrieval, and in ✧ Explainable Style, *Fengyin Lin, Mingkang Li, Da Li, Timothy Hospedales, Yi-Zhe Song, Yonggang Qi*
261. Semantic-Conditional Diffusion Networks for Image Captioning, *Jianjie Luo, Yehao Li, Yingwei Pan, Ting Yao, Jianlin Feng, Hongyang Chao, Tao Mei*
262. REVEAL: Retrieval-Augmented Visual-Language Pre-Training ✧ With Multi-Source Multimodal Knowledge Memory, *Ziniu Hu, Ahmet Iscen, Chen Sun, Zirui Wang, Kai-Wei Chang, Yizhou Sun, Cordelia Schmid, David A. Ross, Alireza Fathi*
263. Variational Distribution Learning for Unsupervised Text-to-Image Generation, *Minsoo Kang, Doyup Lee, Jiseob Kim, Saehoon Kim, Bohyung Han*
264. Scaling Language-Image Pre-Training via Masking, *Yanghao Li, Haoqi Fan, Ronghang Hu, Christoph Feichtenhofer, Kaiming He*
265. LANIT: Language-Driven Image-to-Image Translation for Unlabeled Data, *Jihye Park, Sunwoo Kim, Soohyun Kim, Seokju Cho, Jaejun Yoo, Youngjung Uh, Seungryong Kim*
266. Revisiting Self-Similarity: Structural Embedding for Image Retrieval, *Seongwon Lee, Suhyeon Lee, Hongje Seong, Euntai Kim*
267. Improving Cross-Modal Retrieval With Set of Diverse ✧ Embeddings, *Dongwon Kim, Namyup Kim, Suha Kwak*
268. Masked Autoencoding Does Not Help Natural Language Supervision at Scale, *Floris Weers, Vaishaal Shankar, Angelos Katharopoulos, Yinfei Yang, Tom Gunter*
269. Few-Shot Learning With Visual Distribution Calibration and Cross-Modal Distribution Alignment, *Runqi Wang, Hao Zheng, Xiaoyue Duan, Jianzhuang Liu, Yuning Lu, Tian Wang, Songcen Xu, Baochang Zhang*
270. Deep Hashing With Minimal-Distance-Separated Hash Centers, *Liangdao Wang, Yan Pan, Cong Liu, Hanjiang Lai, Jian Yin, Ye Liu*
271. ConZIC: Controllable Zero-Shot Image Captioning by Sampling-Based Polishing, *Zequn Zeng, Hao Zhang, Ruiying Lu, Dongsheng Wang, Bo Chen, Zhengjue Wang*
272. Learning to Name Classes for Vision and Language Models, *Sarah Parisot, Yongxin Yang, Steven McDonagh*
273. Data-Efficient Large Scale Place Recognition With Graded Similarity Supervision, *María Leyva-Vallina, Nicola Strisciuglio, Nicolai Petkov*
274. DetCLIPv2: Scalable Open-Vocabulary Object Detection Pre-Training via Word-Region Alignment, *Lewei Yao, Jianhua Han, Xiaodan Liang, Dan Xu, Wei Zhang, Zhenguo Li, Hang Xu*
275. HOICLIP: Efficient Knowledge Transfer for HOI Detection With Vision-Language Models, *Shan Ning, Longtian Qiu, Yongfei Liu, Xuming He*
276. OvarNet: Towards Open-Vocabulary Object Attribute Recognition, *Keyan Chen, Xiaolong Jiang, Yao Hu, Xu Tang, Yan Gao, Jianqi Chen, Weidi Xie*
277. NeRF-RPN: A General Framework for Object Detection in NeRFs, *Benran Hu, Junkai Huang, Yichen Liu, Yu-Wing Tai, Chi-Keung Tang*
278. Mask-Free OVIS: Open-Vocabulary Instance Segmentation Without Manual Mask Annotations, *Vibashan VS, Ning Yu, Chen Xing, Can Qin, Mingfei Gao, Juan Carlos Niebles, Vishal M. Patel, Ran Xu*
279. GP-VTON: Towards General Purpose Virtual Try-On via Collaborative Local-Flow Global-Parsing Learning, *Zhenyu Xie, Zaiyu Huang, Xin Dong, Fuwei Zhao, Haoye Dong, Xijin Zhang, Feida Zhu, Xiaodan Liang*
280. Decomposed Soft Prompt Guided Fusion Enhancing for Compositional Zero-Shot Learning, *Xiaocheng Lu, Song Guo, Ziming Liu, Jingcai Guo*
281. Contrastive Grouping With Transformer for Referring Image Segmentation, *Jiajin Tang, Ge Zheng, Cheng Shi, Sibe Yang*
282. GRES: Generalized Referring Expression Segmentation, *Chang ✧ Liu, Henghui Ding, Xudong Jiang*
283. Network-Free, Unsupervised Semantic Segmentation With Synthetic Images, *Qianli Feng, Raghudeep Gadde, Wentong Liao, Eduard Ramon, Aleix Martinez*
284. Few-Shot Semantic Image Synthesis With Class Affinity Transfer, *Marlène Careil, Jakob Verbeek, Stéphane Lathuilière*
285. Ultra-High Resolution Segmentation With Ultra-Rich Context: A Novel Benchmark, *Deyi Ji, Feng Zhao, Hongtao Lu, Mingyuan Tao, Jieping Ye*
286. Content-Aware Token Sharing for Efficient Semantic Segmentation With Vision Transformers, *Chenyang Lu, Daan de Geus, Gijs Dubbelman*
287. Hierarchical Dense Correlation Distillation for Few-Shot ✧ Segmentation, *Bohao Peng, Zhuotao Tian, Xiaoyang Wu, Chengyao Wang, Shu Liu, Jingyong Su, Jiaya Jia*
288. On Calibrating Semantic Segmentation Models: Analyses and an Algorithm, *Dongdong Wang, Boqing Gong, Liqiang Wang*
289. FastInst: A Simple Query-Based Model for Real-Time Instance Segmentation, *Junjie He, Pengyu Li, Yifeng Geng, Xuansong Xie*
290. Out-of-Candidate Rectification for Weakly Supervised Semantic Segmentation, *Zesen Cheng, Pengchong Qiao, Kehan Li, Siheng Li, Pengxu Wei, Xiangyang Ji, Li Yuan, Chang Liu, Jie Chen*
291. Foundation Model Drives Weakly Incremental Learning for Semantic Segmentation, *Chaohui Yu, Qiang Zhou, Jingliang Li, Jianlong Yuan, Zhibin Wang, Fan Wang*
292. Long-Tailed Visual Recognition via Self-Heterogeneous Integration With Knowledge Excavation, *Yan Jin, Mengke Li, Yang Lu, Yiu-ming Cheung, Hanzhi Wang*
293. Instance-Specific and Model-Adaptive Supervision for Semi-Supervised Semantic Segmentation, *Zhen Zhao, Sifan Long, Jimin Pi, Jingdong Wang, Luping Zhou*
294. Active Finetuning: Exploiting Annotation Budget in the Pretraining-Finetuning Paradigm, *Yichen Xie, Han Lu, Junchi Yan, Xiaokang Yang, Masayoshi Tomizuka, Wei Zhan*
295. IDGI: A Framework to Eliminate Explanation Noise From Integrated Gradients, *Ruo Yang, Binghui Wang, Mustafa Bilgic*
296. Weakly Supervised Posture Mining for Fine-Grained Classification, *Zhenchao Tang, Hualin Yang, Calvin Yu-Chian Chen*
297. Vision Transformers Are Good Mask Auto-Labelers, *Shiyi Lan, Xitong Yang, Zhiding Yu, Zuxuan Wu, Jose M. Alvarez, Anima Anandkumar*

298. Enhanced Training of Query-Based Object Detection via Selective Query Recollection, *Fangyi Chen, Han Zhang, Kai Hu, Yu-Kai Huang, Chenchen Zhu, Marios Savvides*
299. Box-Level Active Detection, *Mengyao Lyu, Jundong Zhou, Hui Chen, Yijie Huang, Dongdong Yu, Yaqian Li, Yandong Guo, Yuchen Guo, Liuyu Xiang, Guiguang Ding*
300. CIGAR: Cross-Modality Graph Reasoning for Domain Adaptive Object Detection, *Yabo Liu, Jinghua Wang, Chao Huang, Yaowei Wang, Yong Xu*
301. DA-DETR: Domain Adaptive Detection Transformer With Information Fusion, *Jingyi Zhang, Jiaxing Huang, Zhipeng Luo, Gongjie Zhang, Xiaoqin Zhang, Shijian Lu*
302. Continual Detection Transformer for Incremental Object Detection, *Yaoyao Liu, Bernt Schiele, Andrea Vedaldi, Christian Rupprecht*
303. Semi-DETR: Semi-Supervised Object Detection With Detection Transformers, *Jiacheng Zhang, Xiangru Lin, Wei Zhang, Kuo Wang, Xiao Tan, Junyu Han, Errui Ding, Jingdong Wang, Guanbin Li*
304. Hierarchical Supervision and Shuffle Data Augmentation for 3D Semi-Supervised Object Detection, *Chuangdong Liu, Chenqiang Gao, Fangcen Liu, Pengcheng Li, Deyu Meng, Xinbo Gao*
305. Harmonious Teacher for Cross-Domain Object Detection, *Jinhong Deng, Dongli Xu, Wen Li, Lixin Duan*
306. Contrastive Mean Teacher for Domain Adaptive Object Detectors, *Shengcao Cao, Dhiraj Joshi, Liang-Yan Gui, Yu-Xiong Wang*
307. Out-of-Distributed Semantic Pruning for Robust Semi-Supervised Learning, *Yu Wang, Pengchong Qiao, Chang Liu, Guoli Song, Xiewu Zheng, Jie Chen*
308. (ML)²P-Encoder: On Exploration of Channel-Class Correlation for Multi-Label Zero-Shot Learning, *Ziming Liu, Song Guo, Xiaocheng Lu, Jingcai Guo, Jiawei Zhang, Yue Zeng, Fushuo Huo*
309. MagicNet: Semi-Supervised Multi-Organ Segmentation via Magic-Cube Partition and Recovery, *Duowen Chen, Yunhao Bai, Wei Shen, Qingli Li, Lequan Yu, Yan Wang*
310. Devil Is in the Queries: Advancing Mask Transformers for Real-World Medical Image Segmentation and Out-of-Distribution Localization, *Mingze Yuan, Yingda Xia, Hexin Dong, Zifan Chen, Jiawen Yao, Mingyan Qiu, Ke Yan, Xiaoli Yin, Yu Shi, Xin Chen, Zaiyi Liu, Bin Dong, Jingren Zhou, Le Lu, Ling Zhang, Li Zhang*
311. SQUID: Deep Feature In-Painting for Unsupervised Anomaly Detection, *Tiange Xiang, Yixiao Zhang, Yongyi Lu, Alan L. Yuille, Chaoyi Zhang, Weidong Cai, Zongwei Zhou*
312. OCELOT: Overlapped Cell on Tissue Dataset for Histopathology, *Jeongun Ryu, Aaron Valero Puche, JaeWoong Shin, Seonwook Park, Biagio Brattoli, Jinhee Lee, Wonkyung Jung, Soo Ick Cho, Kyunghyun Paeng, Chan-Young Ock, Donggeun Yoo, Sérgio Pereira*
313. DEGPR: Deep Guided Posterior Regularization for Multi-Class Cell Detection and Counting, *Aayush Kumar Tyagi, Chirag Mohapatra, Prasenjit Das, Govind Makharia, Lalita Mehra, Prathosh AP, Mausam*
314. Best of Both Worlds: Multimodal Contrastive Learning With Tabular and Imaging Data, *Paul Hager, Martin J. Menten, Daniel Rueckert*
315. RankMix: Data Augmentation for Weakly Supervised Learning of Classifying Whole Slide Images With Diverse Sizes and Imbalanced Categories, *Yuan-Chih Chen, Chun-Shien Lu*
316. GEN: Pushing the Limits of Softmax-Based Out-of-Distribution Detection, *Xixi Liu, Yaroslava Lochman, Christopher Zach*
317. Discriminating Known From Unknown Objects via Structure-Enhanced Recurrent Variational AutoEncoder, *Aming Wu, Cheng Deng*
318. Sample-Level Multi-View Graph Clustering, *Yuze Tan, Yixi Liu, Shudong Huang, Wentao Feng, Jiancheng Lv*
319. On the Effects of Self-Supervision and Contrastive Alignment in Deep Multi-View Clustering, *Daniel J. Trosten, Sigurd Løkse, Robert Jenssen, Michael C. Kampffmeyer*
320. Deep Fair Clustering via Maximizing and Minimizing Mutual Information: Theory, Algorithm and Metric, *Pengxin Zeng, Yunfan Li, Peng Hu, Dezhong Peng, Jiancheng Lv, Xi Peng*
321. Transductive Few-Shot Learning With Prototype-Based Label Propagation by Iterative Graph Refinement, *Hao Zhu, Piotr Koniusz*
322. Open-Set Likelihood Maximization for Few-Shot Learning, *Malik Boudiaf, Etienne Bennequin, Myriam Tami, Antoine Toubhans, Pablo Piantanida, Celine Hudelot, Ismail Ben Ayed*
323. HyperMatch: Noise-Tolerant Semi-Supervised Learning via Relaxed Contrastive Constraint, *Beitong Zhou, Jing Lu, Kerui Liu, Yunlu Xu, Zhanzhan Cheng, Yi Niu*
324. Token Boosting for Robust Self-Supervised Visual Transformer Pre-Training, *Tianjiao Li, Lin Geng Foo, Ping Hu, Xindi Shang, Hossein Rahmani, Zehuan Yuan, Jun Liu*
325. Difficulty-Based Sampling for Debaised Contrastive Representation Learning, *Taeuk Jang, Xiaoqian Wang*
326. Improving Selective Visual Question Answering by Learning From Your Peers, *Corentin Dancette, Spencer Whitehead, Rishabh Maheshwary, Ramakrishna Vedantam, Stefan Scherer, Xinlei Chen, Matthieu Cord, Marcus Rohrbach*
327. Superclass Learning With Representation Enhancement, *Zeyu Gan, Suyun Zhao, Jinlong Kang, Liyuan Shang, Hong Chen, Cuiping Li*
328. DISC: Learning From Noisy Labels via Dynamic Instance-Specific Selection and Correction, *Yifan Li, Hu Han, Shiguang Shan, Xilin Chen*
329. FCC: Feature Clusters Compression for Long-Tailed Visual Recognition, *Jian Li, Ziyao Meng, Daqian Shi, Rui Song, Xiaolei Diao, Jingwen Wang, Hao Xu*
330. Dynamically Instance-Guided Adaptation: A Backward-Free Approach for Test-Time Domain Adaptive Semantic Segmentation, *Wei Wang, Zhun Zhong, Weijie Wang, Xi Chen, Charles Ling, Boyu Wang, Nicu Sebe*
331. Semi-Supervised Domain Adaptation With Source Label Adaptation, *Yu-Chu Yu, Hsuan-Tien Lin*
332. Adjustment and Alignment for Unbiased Open Set Domain Adaptation, *Wuyang Li, Jie Liu, Bo Han, Yixuan Yuan*
333. C-SFDA: A Curriculum Learning Aided Self-Training Framework for Efficient Source Free Domain Adaptation, *Nazmul Karim, Niluthpol Chowdhury Mithun, Abhinav Rajvanshi, Han-pang Chiu, Supun Samarasekera, Nazanin Rahnavard*
334. ALOFT: A Lightweight MLP-Like Architecture With Dynamic Low-Frequency Transform for Domain Generalization, *Jintao Guo, Na Wang, Lei Qi, Yinghuan Shi*

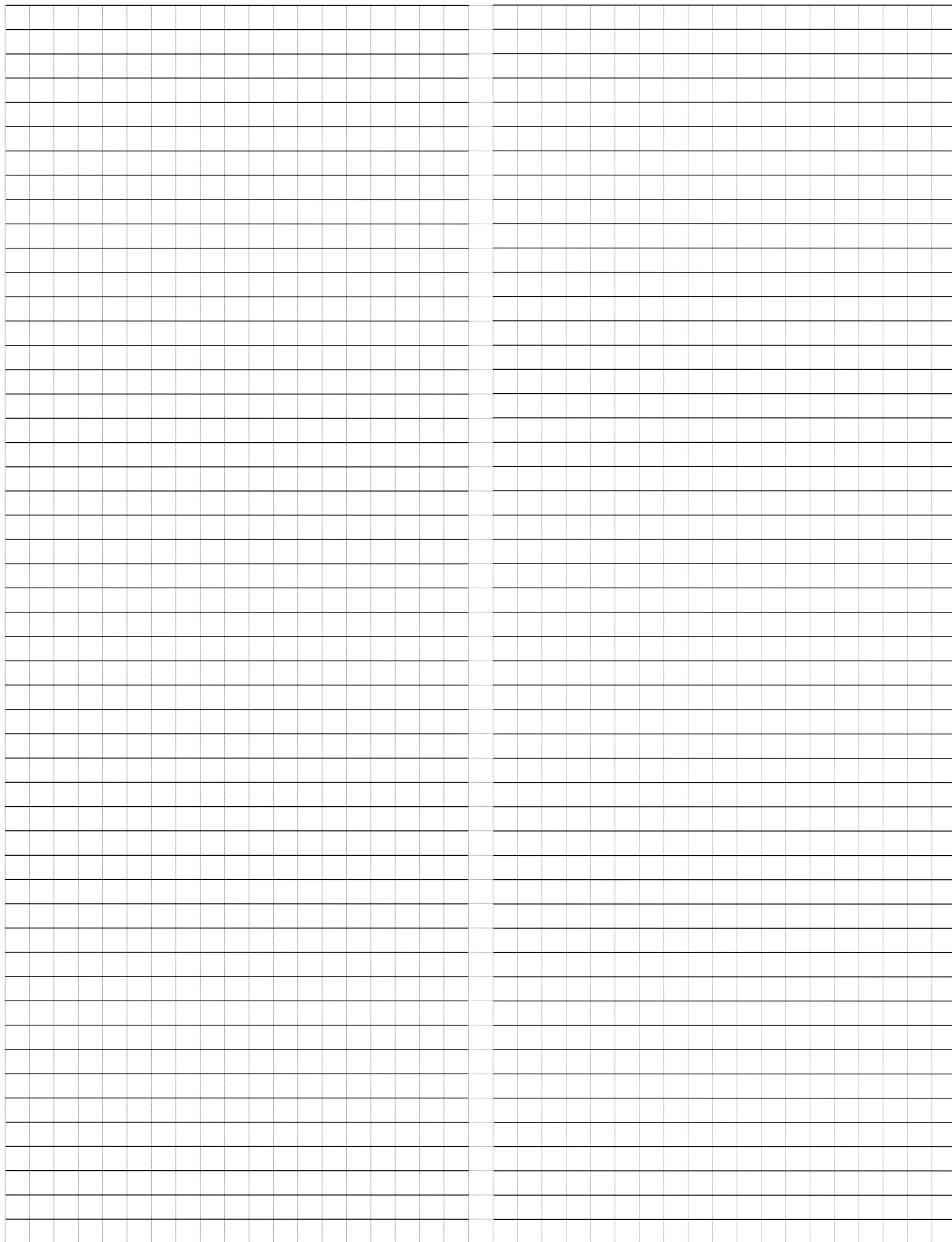
335. Modality-Agnostic Debiasing for Single Domain Generalization, *Sanqing Qu, Yingwei Pan, Guang Chen, Ting Yao, Changjun Jiang, Tao Mei*
336. ActMAD: Activation Matching to Align Distributions for Test-Time-Training, *Muhammad Jehanzeb Mirza, Pol Jané Soneira, Wei Lin, Mateusz Kozinski, Horst Possegger, Horst Bischof*
337. TIPI: Test Time Adaptation With Transformation Invariance, *A. Tuan Nguyen, Thanh Nguyen-Tang, Ser-Nam Lim, Philip H.S. Torr*
338. Improved Test-Time Adaptation for Domain Generalization, *Liang Chen, Yong Zhang, Yibing Song, Ying Shan, Lingqiao Liu*
339. Learning With Fantasy: Semantic-Aware Virtual Contrastive Constraint for Few-Shot Class-Incremental Learning, *Zeyin Song, Yifan Zhao, Yujun Shi, Peixi Peng, Li Yuan, Yonghong Tian*
340. NIFF: Alleviating Forgetting in Generalized Few-Shot Object Detection via Neural Instance Feature Forging, *Karim Guirguis, Johannes Meier, George Eskandar, Matthias Kayser, Bin Yang, Jürgen Beyerer*
341. MixPHM: Redundancy-Aware Parameter-Efficient Tuning for Low-Resource Visual Question Answering, *Jingjing Jiang, Nanning Zheng*
342. PIVOT: Prompting for Video Continual Learning, *Andrés Villa, Juan León Alcázar, Motasem Alfarra, Kumail Alhamoud, Julio Hurtado, Fabian Caba Heilbron, Alvaro Soto, Bernard Ghanem*
343. BlackVIP: Black-Box Visual Prompting for Robust Transfer Learning, *Changdae Oh, Hyeji Hwang, Hee-young Lee, YongTaek Lim, Geunyoung Jung, Jiyoung Jung, Hosik Choi, Kyungwoo Song*
344. DKT: Diverse Knowledge Transfer Transformer for Class Incremental Learning, *Xinyuan Gao, Yuhang He, Songlin Dong, Jie Cheng, Xing Wei, Yihong Gong*
345. PCR: Proxy-Based Contrastive Replay for Online Class-Incremental Continual Learning, *Huiwei Lin, Baoquan Zhang, Shanshan Feng, Xutao Li, Yunming Ye*
346. Masked Autoencoders Enable Efficient Knowledge Distillers, *Yutong Bai, Zeyu Wang, Junfei Xiao, Chen Wei, Huiyu Wang, Alan L. Yuille, Yuyin Zhou, Cihang Xie*
347. Data-Free Knowledge Distillation via Feature Exchange and Activation Region Constraint, *Shikang Yu, Jiachen Chen, Hu Han, Shuqiang Jiang*
348. Multi-Level Logit Distillation, *Ying Jin, Jiaqi Wang, Dahua Lin*
349. Preserving Linear Separability in Continual Learning by Backward Feature Projection, *Qiao Gu, Dongsub Shim, Florian Shkurti*
350. Critical Learning Periods for Multisensory Integration in Deep Networks, *Michael Kleinman, Alessandro Achille, Stefano Soatto*
351. SLACK: Stable Learning of Augmentations With Cold-Start and KL Regularization, *Juliette Marrie, Michael Arbel, Diane Larlus, Julien Mairal*
352. Improving Generalization With Domain Convex Game, *Fangrui Lv, Jian Liang, Shuang Li, Jinming Zhang, Di Liu*
353. Exploring Data Geometry for Continual Learning, *Zhi Gao, Chen Xu, Feng Li, Yunde Jia, Mehrtash Harandi, Yuwei Wu*
354. FlowGrad: Controlling the Output of Generative ODEs With Gradients, *Xingchao Liu, Lemeng Wu, Shujian Zhang, Chengyue Gong, Wei Ping, Qiang Liu*
355. Deep Graph Reprogramming, *Yongcheng Jing, Chongbin Yuan, Li Ju, Yiding Yang, Xinchao Wang, Dacheng Tao*
356. X-Pruner: eXplainable Pruning for Vision Transformers, *Lu Yu, Wei Xiang*
357. Bias in Pruned Vision Models: In-Depth Analysis and Countermeasures, *Eugenia Iofinova, Alexandra Peste, Dan Alistarh*
358. Compacting Binary Neural Networks by Sparse Kernel Selection, *Yikai Wang, Wenbing Huang, Yinpeng Dong, Fuchun Sun, Anbang Yao*
359. Deep Deterministic Uncertainty: A New Simple Baseline, *Jishnu Mukhoti, Andreas Kirsch, Joost van Amersfoort, Philip H.S. Torr, Yarin Gal*
360. Understanding Deep Generative Models With Generalized Empirical Likelihoods, *Suman Ravuri, Mélanie Rey, Shakir Mohamed, Marc Peter Deisenroth*
361. Fair Scratch Tickets: Finding Fair Sparse Networks Without Weight Training, *Pengwei Tang, Wei Yao, Zhicong Li, Yong Liu*
362. Hard Sample Matters a Lot in Zero-Shot Quantization, *Huantong Li, Xiangmiao Wu, Fanbing Lv, Daihai Liao, Thomas H. Li, Yonggang Zhang, Bo Han, Mingkui Tan*
363. PD-Quant: Post-Training Quantization Based on Prediction Difference Metric, *Jiawei Liu, Lin Niu, Zhihang Yuan, Dawei Yang, Xinggang Wang, Wenyu Liu*
364. Vector Quantization With Self-Attention for Quality-Independent Representation Learning, *Zhou Yang, Weisheng Dong, Xin Li, Mengluan Huang, Yulin Sun, Guangming Shi*
365. Masked Auto-Encoders Meet Generative Adversarial Networks and Beyond, *Zhengcong Fei, Mingyuan Fan, Li Zhu, Junshi Huang, Xiaoming Wei, Xiaolin Wei*
366. Sequential Training of GANs Against GAN-Classifiers Reveals Correlated “Knowledge Gaps” Present Among Independently Trained GAN Instances, *Arkanath Pathak, Nicholas Dufour*
367. Edges to Shapes to Concepts: Adversarial Augmentation for Robust Vision, *Aditay Tripathi, Rishubh Singh, Anirban Chakraborty, Pradeep Shenoy*
368. Towards Universal Fake Image Detectors That Generalize Across Generative Models, *Utkarsh Ojha, Yuheng Li, Yong Jae Lee*
369. Explicit Boundary Guided Semi-Push-Pull Contrastive Learning for Supervised Anomaly Detection, *Xincheng Yao, Ruqi Li, Jing Zhang, Jun Sun, Chongyang Zhang*
370. Generating Anomalies for Video Anomaly Detection With Prompt-Based Feature Mapping, *Zuhao Liu, Xiao-Ming Wu, Dian Zheng, Kun-Yu Lin, Wei-Shi Zheng*
371. Revisiting Reverse Distillation for Anomaly Detection, *Tran Dinh Tien, Anh Tuan Nguyen, Nguyen Hoang Tran, Ta Duc Huy, Soan T.M. Duong, Chanh D. Tr. Nguyen, Steven Q. H. Truong*
372. MetaMix: Towards Corruption-Robust Continual Learning With Temporally Self-Adaptive Data Transformation, *Zhenyi Wang, Li Shen, Donglin Zhan, Qiuling Suo, Yanjun Zhu, Tiehang Duan, Mingchen Gao*
373. ScaleFL: Resource-Adaptive Federated Learning With Heterogeneous Clients, *Fatih Ilhan, Gong Su, Ling Liu*
374. Confidence-Aware Personalized Federated Learning via Variational Expectation Maximization, *Junyi Zhu, Xingchen Ma, Matthew B. Blaschko*
375. Make Landscape Flatter in Differentially Private Federated Learning, *Yifan Shi, Yingqi Liu, Kang Wei, Li Shen, Xueqian Wang, Dacheng Tao*

376. Rethinking Domain Generalization for Face Anti-Spoofing: Separability and Alignment, *Yiyao Sun, Yaojie Liu, Xiaoming Liu, Yixuan Li, Wen-Sheng Chu*
377. StyleAdv: Meta Style Adversarial Training for Cross-Domain Few-Shot Learning, *Yuqian Fu, Yu Xie, Yanwei Fu, Yu-Gang Jiang*
378. The Dazrk Side of Dynamic Routing Neural Networks: Towards Efficiency Backdoor Injection, *Simin Chen, Hanlin Chen, Mirazul Haque, Cong Liu, Wei Yang*
379. Architectural Backdoors in Neural Networks, *Mikel Bober-Irizar, Iliia Shumailov, Yiren Zhao, Robert Mullins, Nicolas Papernot*
380. You Are Catching My Attention: Are Vision Transformers Bad Learners Under Backdoor Attacks? *Zenghui Yuan, Pan Zhou, Kai Zou, Yu Cheng*
381. A Practical Upper Bound for the Worst-Case Attribution Deviations, *Fan Wang, Adams Wai-Kin Kong*
382. Sibling-Attack: Rethinking Transferable Adversarial Attacks Against Face Recognition, *Zexin Li, Bangjie Yin, Taiping Yao, Junfeng Guo, Shouhong Ding, Simin Chen, Cong Liu*
383. Angelic Patches for Improving Third-Party Object Detector Performance, *Wenwen Si, Shuo Li, Sangdon Park, Insup Lee, Osbert Bastani*
384. Introducing Competition to Boost the Transferability of Targeted Adversarial Examples Through Clean Feature Mixup, *Junyoung Byun, Myung-Joon Kwon, Seungju Cho, Yoonji Kim, Changick Kim*
385. Towards Compositional Adversarial Robustness: Generalizing Adversarial Training to Composite Semantic Perturbations, *Lei Hsiung, Yun-Yun Tsai, Pin-Yu Chen, Tsung-Yi Ho*
386. Boosting Accuracy and Robustness of Student Models via Adaptive Adversarial Distillation, *Bo Huang, Mingyang Chen, Yi Wang, Junda Lu, Minhao Cheng, Wei Wang*
387. The Enemy of My Enemy Is My Friend: Exploring Inverse Adversaries for Improving Adversarial Training, *Junhao Dong, Seyed-Mohsen Moosavi-Dezfooli, Jianhuang Lai, Xiaohua Xie*
388. Robust Single Image Reflection Removal Against Adversarial Attacks, *Zhenbo Song, Zhenyuan Zhang, Kaihao Zhang, Wenhan Luo, Zhaoxin Fan, Wenqi Ren, Jianfeng Lu*
389. Physical-World Optical Adversarial Attacks on 3D Face Recognition, *Yanjie Li, Yiquan Li, Xuelong Dai, Songtao Guo, Bin Xiao*
390. AUNet: Learning Relations Between Action Units for Face Forgery Detection, *Weiming Bai, Yufan Liu, Zhipeng Zhang, Bing Li, Weiming Hu*

AI Art Gallery Participants

We are pleased to recognize the following participants in the CVPR 2023 AI Art Gallery. The AI Art Gallery is displayed in the CVPR 2023 Demo area as well as other locations — showcasing AI-powered or generative artistic works or other visually engaging results produced through computer vision research.

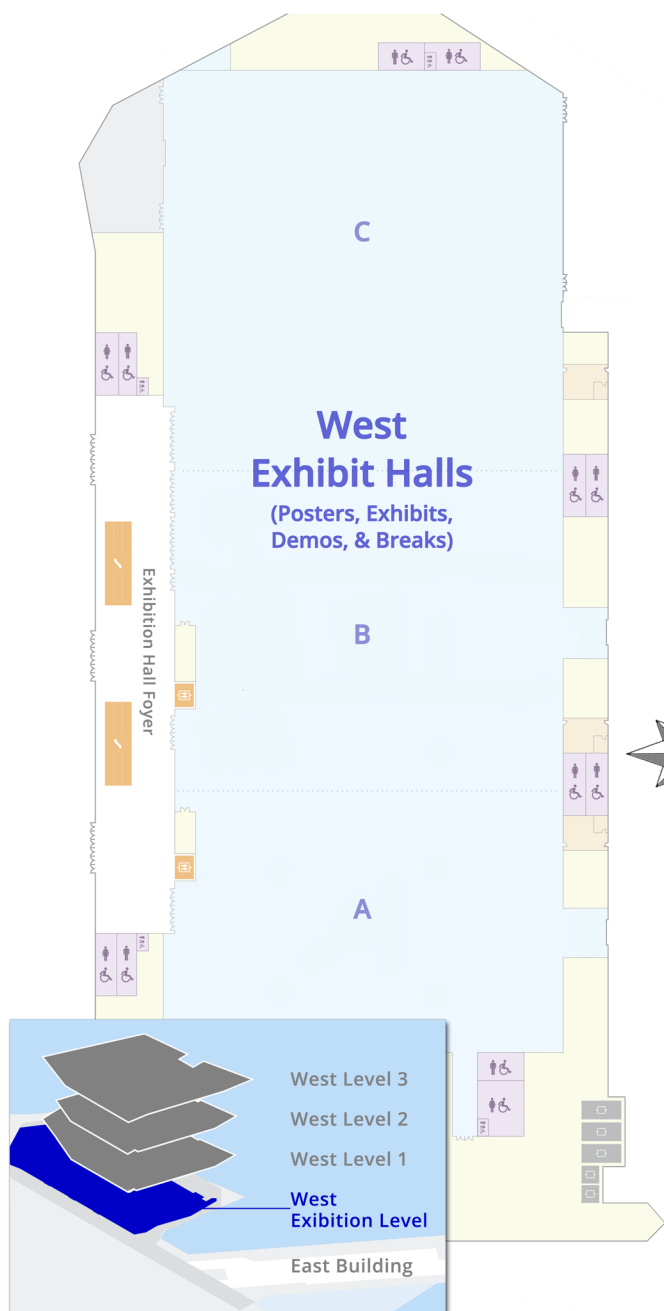
Abhishek Das	Corey Karnei	James E Perlman	Nicholas Dutka	Vasudev Lal
Addin Cui	Cyrus Vachha	Jarno Tuominen	Nirvaal	William Stanford
AiDA	Danny Cisco	Jean Yu	Omri Avrahami	Xiao Ma (Maximal Margin)
Alan Bovik	David Marx	Jingkun An	Pavan Turaga	Xijun Wang
Alex Redden	Devanshu Tak	Jonathan Stephens	Pengjiang	Xinlei Bao
Alice Zou	Devi Parikh	Kayla	Prakash Perumal Haridas	Xinyuan Wang
Andreea Pocol	Diana Wofk	Kelly Xu	Riccardo Bazzoni	Xuekun Jiang
Ang Li	Diego Porres	Kim Hea Yeon	Richard Ely Locke Higgins	Yang Ou
Anush	Enshen Zhou	Kyle Goodrich	RJ's RATDREAMS	Yejin Kim
Anyi Rao	Estelle Aflalo	Lesley Istead	Rui-Qi Zhang	Yewon Kim
Arjan van Asperen	Fabio Nonato	Lior Azi	Sajith Rajapaksa	Yi-Xiao Li
Arshia Sobhan	Fikret Ukis	Lodewijk Luijt	Samuel Pietri	Yinghao Zhu
Bilawal Sidhu	Franc Lucent	Long He	Sanghoon Park	Yiran Wang
Bo Dai	Gabriela Ben Melech	Luc Weytingh	Scottie Fox	Yu Liu
Bo Hou	Grade Eterna	Made by Saurus	Shao-Yen Tseng	Yu Zhu
Byoung-Ki Jeon	Hannah Johnston	Magzhan Gabidolla	Sherman Siu	Yuan-Fu Yang
Caixin Kang	Hao Li	Matteo Castaldo	Stefan Babic	Yuki Asano
CC Laan	Haoran Feng	Matthias Müller	Stefanie Zollmann	Yuqian Sun
Chenfei Wu	Hugues Bruyère	Megha Srivastava	Subin Kim	Yuwei Guo
Christian Limberg	Ian George	Mehow	Sunhee Hwang	Zhujun (Grace) Yao
Chuang Wang	Infinite Vibes	Miguel Carreira-Perpinan	Sunny Sanyal	Zihao Li
Clinton Wang	Iuan-Kai Fang	Moshe Mandel	Symbiotic/Symctrae	Zihou Ng
CoffeeVectors	Jake Oleson	Nan Duan	Uttam Grandhi	Ziyi Wang



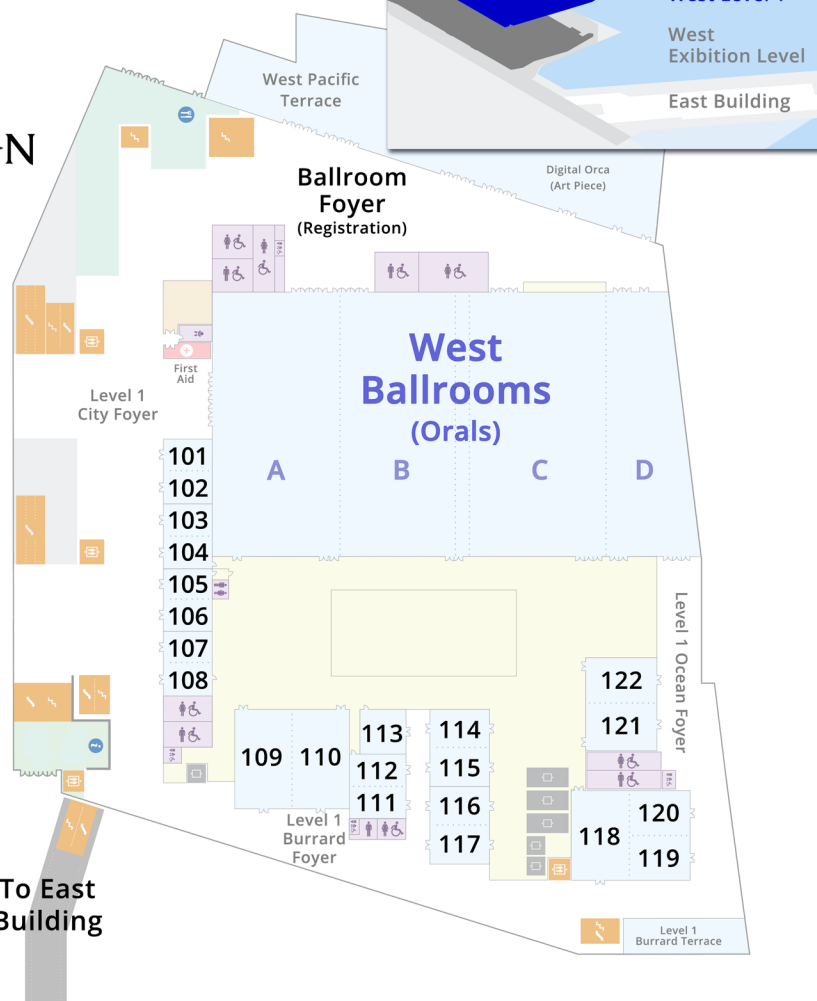
JUNE 18-22, 2023

CVPR VANCOUVER, CANADA

VANCOUVER
CONVENTION
CENTRE



See opposite page for:
West Level 3 and
East Building maps



CVPR 2023 Sponsors



IEEE
COMPUTER
SOCIETY



JUNE 18-22, 2023

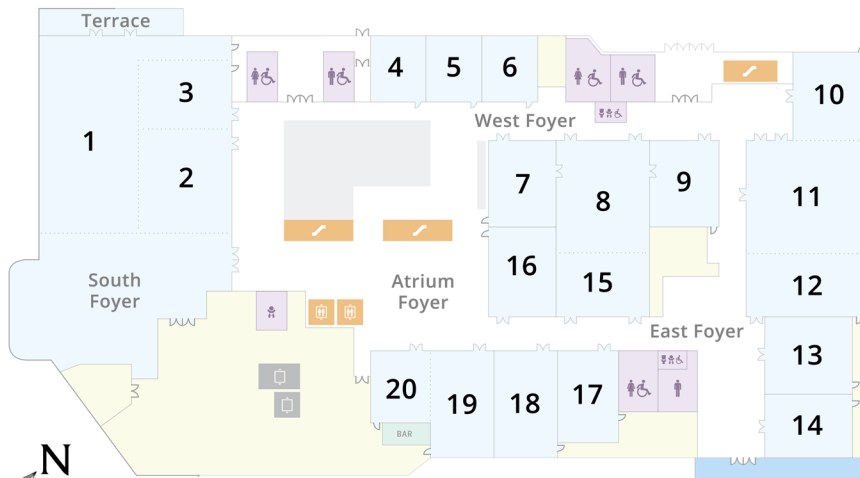
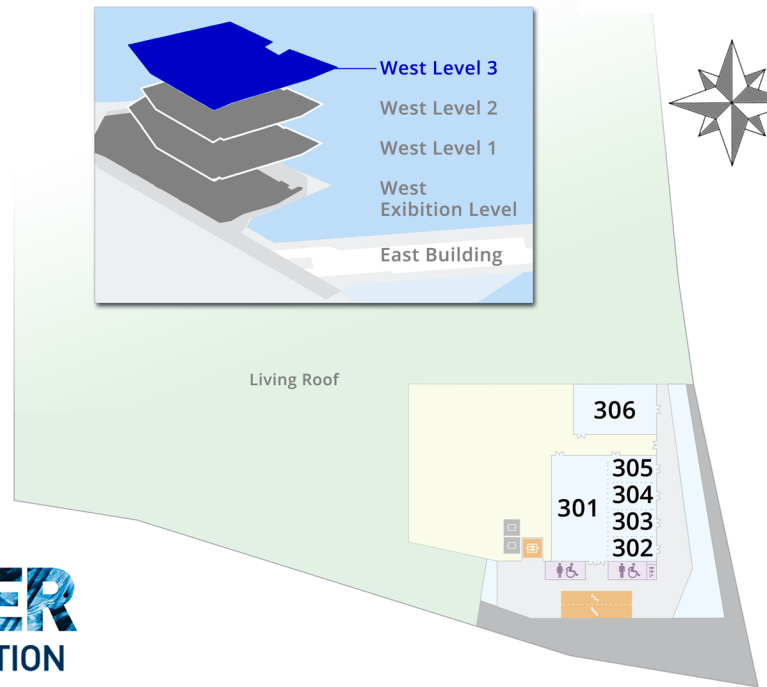
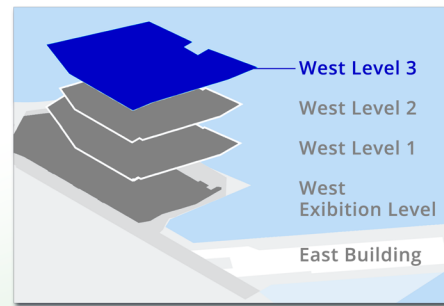
CVPR



CVPR 2023 (slido):
For Q&A, polls, etc.
during plenary events.



VANCOUVER CONVENTION CENTRE



See opposite page for:
West Exhibit Level,
West Level 1, and
West Level 2 maps

