



IEEE/CVF Conference on
**Computer Vision and
Pattern Recognition**

Program Guide

CVPR
June 16 – 20, 2019
Long Beach, California

CVPR 2019 ORGANIZING COMMITTEE

General Chairs



Larry Davis
University of Maryland



Philip Torr
University of Oxford



Song-Chun Zhu
University of California,
Los Angeles

Program Chairs



Abhinav Gupta
Carnegie Mellon
University



Derek Hoiem
University of Illinois
at Urbana-Champaign



Gang Hua
Wormpex.AI Research



Zhuowen Tu
University of California,
San Diego

AC Meeting Chairs



Manmohan Chandraker
University of California,
San Diego



Hao Su
University of California,
San Diego

Workshop Chairs



Sanja Fidler
University of Toronto



Andrea Vedaldi
University of Oxford

Tutorials Chairs



Ali Farhadi
University of Washington



M. Alex O. Vasilescu
University of California,
Los Angeles

Finance Chairs



Brandon Rothrock
Jet Propulsion Laboratory



Walter Scheirer
University of Notre Dame

Publications Chairs



William Brendel
Snap Inc.



Mohamed R. Amer
RobustAI

Corporate Relations Chairs



Mei Han
Ping An Technology

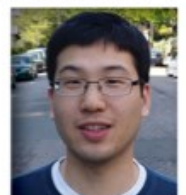


Florent Perronnin
Naver Labs Europe

Demo & Exhibition Chairs



Andrew D. Bagdanov
Università degli Studi di Firenze



Yong Jae Lee
University of California, Davis

Doctoral Consortium Chairs



Olga Russakovsky
Princeton University



Nathan Jacobs
University of Kentucky

Presentation Chairs



Maria Zontak
Microsoft



Tianfu Wu
NC State University

Local Arrangements Chair



Le Dong
DMAI

Registration Chair and Event Planner



Nicole Bumpus Finn
C to C Events

Website Chair

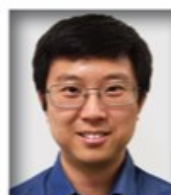


Ming-Ming Cheng
Nankai University

Webmasters



Nishant Shukla
University of California,
Los Angeles



Yixin Zhu
University of California,
Los Angeles



**LONG BEACH
CALIFORNIA
June 16-20, 2019**

Message from the General and Program Chairs

Welcome to the 32nd meeting of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2019) at Long Beach, CA. CVPR continues to be one of the best venues for researchers in our community to present their most exciting advances in computer vision, pattern recognition, machine learning, robotics, and artificial intelligence. With oral and poster presentations, tutorials, workshops, demos, an ever-growing number of exhibitions, and numerous social events, this is a week that everyone will enjoy. The co-location with ICML 2019 this year brings opportunity for more exchanges between the communities.

Accelerating a multi-year trend, CVPR continued its rapid growth with a record number of 5165 submissions. After three months' diligent work from the 132 area chairs and 2887 reviewers, including those reviewers who kindly helped in the last minutes serving as emergency reviewers, 1300 papers were accepted, yielding 1294 papers in the final program after a few withdrawals. Each paper received at least 3 full reviews, and the acceptance decisions were made within AC triplets in consultation with larger conflict-free panels. The oral/poster decision was made within panels of 12-15 ACs. Following the best practice of our community, the program chairs did not place any restrictions on acceptance. The final acceptance rate is 25.1%, consistent with the high standard of past CVPR conferences.

Out of these 1294 papers, 288 are selected for oral presentations, and all papers have poster presentations. This year, the oral presentations are short, 6 minutes each including transition/questions, so that we may accommodate more oral presentations. Per PAMI-TC policy, program chairs did not submit

papers, which allowed them to be free of conflict in the paper review process. To accommodate the growing number of papers and attendees while maintaining the three-day length of the conference, we run oral sessions in three parallel tracks and devote the entire technical program to accepted papers.

We would like to thank everyone involved in making CVPR 2019 a success. This includes the organizing committee, the area chairs, the reviewers, authors, demo session participants, donors, exhibitors, and everyone else without whom this meeting would not be possible. The program chairs particularly thank a few unsung heroes that helped us tremendously: Eric Mortensen for mentoring the publication chairs and managing camera-ready and program efforts; Ming-Ming Cheng for quickly updating the website; Hao Su for working overtime as both area chair and AC meeting chair; Walter Scheirer for managing finances and serving as the true memory of the organization process; and the Microsoft CMT support team for the tremendous help with prompt responses. We also thank Nicole Finn, Liz Ryan, and C to C Events for their organization of the logistics of the conference.

Last but not least, we thank all of you for attending CVPR 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 the gorgeous Long Beach and great Los Angeles before or after the conference. Enjoy CVPR 2019!!

Program Chairs: Abhinav Gupta, Derek Hoiem,
Gang Hua, and Zhuowen Tu

General Chairs: Larry Davis, Philip Torr, and Song-Chun Zhu

CVPR 2019 Area Chairs

Lourdes Agapito	Stephen Gould	Ajay Kumar	Srinivasa Narasimhan	Deqing Sun
Zeynep Akata	Kristen Grauman	In So Kweon	Juan Carlos Niebles	Jian Sun
Pablo Arbelaez	Saurabh Gupta	Christoph Lampert	Aude Oliva	Ping Tan
Xiang Bai	Gregory D. Hager	Ivan Laptev	Long Quan	Camillo Taylor
Jonathan Barron	Bharath Hariharan	Svetlana Lazebnik	Ravi Ramamoorthi	YingLi Tian
Serge Belongie	Tal Hassner	Erik Learned-Miller	Deva Ramanan	Sinisa Todorovic
Tamara Berg	James Hays	Yong Jae Lee	Nalini Ratha	Antonio Torralba
Horst Bischof	Martial Hebert	Hongdong Li	James Rehg	Matthew Turk
Octavia Camps	Aaron Hertzmann	Yin Li	Xiaofeng Ren	Tinne Tuytelaars
Rama Chellappa	Judy Hoffman	Dahua Lin	Marcus Rohrbach	Anton Van Den Hengel
Xilin Chen	Anthony Hoogs	Liang Lin	Stefan Roth	Laurens van der Maaten
Ming-Ming Cheng	Jia-Bin Huang	Zhe Lin	Olga Russakovsky	Carl Vondrick
Jason Corso	Qixing Huang	Zhouchen Lin	Bryan Russell	David Wipf
Trevor Darrell	Phillip Isola	Haibin Ling	Michael Ryoo	Ying Wu
Jia Deng	Nathan Jacobs	Ce Liu	Dimitris Samaras	Ying Nian Wu
Minh Do	Herve Jegou	Xiaobai Liu	Sudeep Sarkar	Shuicheng Yan
Piotr Dollar	Jiaya Jia	Xiaoming Liu	Harpreet Sawhney	Jingyi Yu
Alexei Efros	Hailin Jin	Yanxi Liu	Alexander Schwing	Stella Yu
Ali Farhadi	Frederic Jurie	Zicheng Liu	Nicu Sebe	Junsong Yuan
Ryan Farrell	Sing Bing Kang	Chen Change Loy	Thomas Serre	Lu Yuan
Cornelia Fermuller	Ira Kemelmacher-Shlizerman	Le Lu	Fei Sha	Ersin Yumer
Andrew Fitzgibbon	Vladlen Koltun	Simon Lucey	Shiguang Shan	Zhengyou Zhang
David Forsyth	Sanjeev Koppal	Jiebo Luo	Abhinav Shrivastava	Qi Zhao
David Fouhey	Jana Kosecka	Yi Ma	Leonid Sigal	S. Kevin Zhou
Jan-Michael Frahm	Adriana Kovashka	Subhransu Maji	Noah Snavely	
Bill Freeman	Philipp Kraehenbuehl	Aleix Martinez	Hao Su	
Georgia Gkioxari	David Kriegman	Philippos Mordohai	Erik Sudderth	

CVPR 2019 Outstanding Reviewers

We are pleased to recognize the following researchers as "CVPR 2019 Outstanding Reviewers". These reviewers contributed at least two reviews noted as excellent by area chairs. The names in bold with asterisks deserve special mention as contributing at least four reviews noted as excellent by area chairs.

Yagiz Aksoy	Heng Fan*	Alexander Kirillov	Vladimir Pavlovic	Maxim Tatarchenko
Samuel Albanie	Christoph Feichtenhofer	Patrick Knöbelreiter	Juan Perez	Damien Teney
Peter Anderson*	Qianli Feng	Reinhard Koch	Loic Peter	Christopher Thomas
Misha Andriluka	Wolfgang Foerstner	Naejin Kong	Michael Pfeiffer	Joseph Tighe
Relja Arandjelović	Jean-Sebastien Franco	Ender Konukoglu	Hanspeter Pfister	Anh Tran
Mathieu Aubry*	Friedrich Fraundorfer	Satwik Kottur	Roman Pflugfelder	Alejandro Troccoli
Vineeth N. Balasubramanian	Silvano Galliani	Iro Laina	Aleksis Pirinen	Shubham Tulsiani
Aayush Bansal	Feng Gao	Jean-Francois Lalonde	Bryan Plummer	Hsiao-Yu Tung
Linchao Bao	Jin Gao	Rynson Lau	Matteo Poggi	Pavan Turaga
Fabian Benitez-Quiroz	Alberto Garcia-Garcia	Stefan Lee	Gerard Pons-Moll	Ambrih Tyagi
Florian Bernard	Jochen Gast	Chen-Yu Lee	Jordi Pont-Tuset*	Jasper Uijlings
Gedas Bertasius	Michael Gharbi	Victor Lempitsky	Ameya Prabhu	Osman Ulusoy
Adel Bibi	Soumya Ghosh*	Gil Levi	Brian Price	Christoph Vogel
Stan Birchfield	Ross Girshick	Jose Lezama*	Jerry Prince	Zhangyang Wang
Yochai Blau	Ioannis Gkioulekas*	Jia Li	Senthil Purushwalkam	Lei Wang
Yuri Boykov	Michael Goesele	Wen Li	Filip Radenovic	Dong Wang
Eric Brachmann	Yash Goyal	Siyang Li	Ilija Radosavovic*	Xintao Wang
Samarth Brahmbhatt	Thibault Groueix	Yujia Li	Rene Ranftl	Yufei Wang
Toby Breckon	Josechu Guerrero	Shengcai Liao	Maheen Rashid	Ruiping Wang*
Sam Buchanan*	Fatma Güney*	Changsong Liu	Emonet Rémi	Anne Wannenwetsch
Sergi Caelles	Qi Guo	Xihui Liu	Emanuele Rodola	Maggie Wigness
Nathan Cahill	Agrim Gupta	Yang Liu	Giorgio Roffo	Chris Williams
Zhaowei Cai	Stefan Haller	Zhaoyang Lv	Anna Rohrbach	Christian Wolf
Zhangjie Cao	Xintong Han	Chao Ma	Aruni RoyChowdhury	Lior Wolf
Ayan Chakrabarti	Adam Harrison*	Oisin Mac Aodha	Shunta Saito	Chao-Yuan Wu
Rudrasis Chakraborty	David Harwath	Michael Maire	Benjamin Sapp	Zhirong Wu
Soravit Changpinyo	Lisa Anne Hendricks	Mateusz Malinowski	Torsten Sattler	Tong Xiao
Siddhartha Chaudhuri	Yannick Hold-Geoffroy	Clement Mallet	Hanno Schar	Jun Xu
Jianbo Chen	Junhui Hou	Massimiliano Mancini	Daniel Scharstein	Qianqian Xu
Yung-Yu Chuang	Zhe Hu	Kevis-Kokitsi Maninis	Walter Scheirer*	Tsun-Yi Yang
Michael Cogswell*	Ahmad Humayun	Edgar Margffoy-Tuay	Konrad Schindler	Xiaodong Yang
Maxwell Collins*	Junhwa Hur	Renaud Marlet*	Sohil Shah	Kwang Moo Yi
John Collomosse	Varun Jampani	Iacopo Masi*	Roman Shapovalov	Lijun Yin
David Crandall*	Joel Janai	Diana Mateus	Gaurav Sharma	Chong You
Dengxin Dai	Suren Jayasuriya	Lars Mescheder	Viktoria Sharmanska	Ke Yu
Adrian Dalca	Simon Jenni*	Liang Mi	Evan Shelhamer	Lei Zhang
Abir Das	Huajie Jiang	Tomer Michaeli	Xiaohui Shen	Richard Zhang
Andrew Davison	Huaizu Jiang	Krystian Mikolajczyk	Miaoqing Shi	Qijun Zhao
Konstantinos Derpanis	Dakai Jin	Hossein Mobahi	Martin Simonovsky	Liang Zheng
Ferran Diego	Kushal Kafle	Michael Moeller	Sudipta Sinha*	Pan Zhou
Thanh-Toan Do	Ioannis Kakadiaris	Roosbeh Mottaghi	Cees Snoek	Xingyi Zhou
Alexey Dosovitskiy	Angjoo Kanazawa*	Vittorio Murino	Yibing Song	Xizhou Zhu
James Duncan	Melih Kandemir	Tushar Nagarajan	Pratul Srinivasan	Jun-Yan Zhu
Chi Nhan Duong*	Misha Kazhdan	Natalia Neverova	Rainer Stiefelhagen	Yixin Zhu
Debidatta Dwibedi	Aniruddha Kembhavi	Benjamin Nuernberger	Jin Sun	Andrew Zisserman*
Noha El-Zehiry	Alex Kendall	Bjorn Ommer	Ju Sun	Maria Zontak
Sergio Escalera	Cem Keskin*	Mohamed Omran	Supasorn Suwajanakorn	
Carlos Esteves	Salman Khan	Jose Oramas	Duy-Nguyen Ta	
Haoqi Fan	Seon Joo Kim	Dipan Pal	Yuichi Taguchi	
Deng-Ping Fan*	Tae-Kyun Kim	Genevieve Patterson	Makarand Tapaswi	

CVPR 2019 Emergency Reviewers

We also want to recognize the following researchers as “CVPR 2019 Emergency Reviewers”. These reviewers were willing to provide an “emergency” review on short notice within a very short time frame. Thank you for your service.

Sathyanarayanan Aakur	Xuanyi Dong	Vicky Kalogeiton	Prakhar Mehrotra	Xiaobo Shen	Qi Wu
Abdelrahman Abdelhamed	Simon Donné	Hakki Karaimer	Jingjing Meng	Ying Shen	Cihang Xie
Varun Agrawal	Alexey Dosovitskiy	Svebor Karaman	Thomas Mensink	Honghui Shi	Saining Xie
Karim Ahmed	Bertram Drost	Nikolaos Karianakis	Liang Mi	Kevin Shih	Wuyuan Xie
Unaiza Ahsan	Yueqi Duan	Parneet Kaur	Pedro Miraldo	Daeyun Shin	Junliang Xing
Naveed Akhtar	Thibaut Durand	Parmeshwar Khurd	Niluthpol Mithun	Zhixin Shu	Lin Xiong
Derya Akkaynak	Debidatta Dwibedi	Pyojin Kim	Pietro Morerio	Suriya Singh	Hongyu Xu
Humam Alwassel	Jan Ernst	Seungryong Kim	Franziska Mueller	Gregory Slabaugh	Jia Xu
Alexander Andreopoulos	Bin Fan	Akisato Kimura	Armin Mustafa	Brandon Smith	Jun Xu
Rushil Anirudh	Martin Fergie	Hedvig Kjellström	Siva Karthik Mustikovela	Patrick Snape	Mingze Xu
Michel Antunes	Victor Fragoso	Laurent Kneip	Sobhan Naderi Parizi	Francesco Solera	Peng Xu
Mathieu Aubry	Christopher Funk	Piotr Koniusz	Shah Nawaz	Dongjin Song	Yanyu Xu
Samaneh Azadi	Raghudeep Gadde	Adam Kortylewski	Natalia Neverova	Yale Song	Zheng Xu
Min Bai	Silvano Galliani	Gurunandan Krishnan	Phuc Nguyen	Yibing Song	Qingan Yan
Ankan Bansal	Junbin Gao	Hilde Kuehne	Shohei Nobuhara	Pratul Srinivasan	Dawei Yang
Lorenzo Baraldi	Yue Gao	Andreas Kuhn	Ferda Ofli	Joerg Stueckler	Ming Yang
Kobus Barnard	Kirill Gavriluk	Arjan Kuijper	Seong Joon Oh	Jong-Chyi Su	Yanchao Yang
Binod Bhattarai	Shiming Ge	Vijay Kumar B G	Mohamed Omran	Yao Sui	Yang Yang
Tolga Birdal	Andrew Gilbert	Iro Laina	Aljosa Osep	Jin Sun	Chengxi Ye
Federica Bogo	Lluis Gomez	Xiangyuan Lan	Martin R. Oswald	Qianru Sun	Jinwei Ye
Terrance Boulton	Priya Goyal	Hieu Le	Matthew O'Toole	Shao-Hua Sun	Mang YE
Eric Brachmann	Li Guan	Huu Le	Dipan Pal	Zhun Sun	Renjiao Yi
Adrian Bulat	Jie Gui	Kwonjoon Lee	Xingang Pan	Chaowei Tan	Xi Yin
Giedrius Burachas	Guodong Guo	Minsik Lee	Nikolaos Passalis	Qingyang Tan	Wong Yongkang
Zoya Bylinskii	Kaiwen Guo	Gil Levi	Geneviève Patterson	Wei Tang	Gang Yu
Narayanan C Krishnan	Tiantong Guo	Evgeny Levinkov	Danda Pani Paudel	Youbao Tang	Jiahui Yu
Nathan Cahill	Isma Hadji	Kai Li	Yuxin Peng	Yuxing Tang	Shouo-I Yu
Qingxing Cao	Xintong Han	Peiyi Li	Tomas Pfister	Makarand Tapaswi	Tao Yu
Zhangjie Cao	Soren Hauberg	Qing Li	AJ Piergiovanni	Ayush Tewari	Huangying Zhan
Luca Carlone	Zeeshan Hayder	Shaozi Li	Matteo Poggi	Christopher Thomas	Bowen Zhang
Hakan Cevikalp	Junfeng He	Shiwei Li	Jordi Pont-Tuset	Kinh Tieu	Chi Zhang
Rudrasis Chakraborty	Lei He	Paul Pu Liang	Alin-Ionut Popa	Chetan Tonde	Jianming Zhang
Shayok Chakraborty	Lifang He	Minghui Liao	Omid Poursaeed	Wei-Chih Tu	Li Zhang
Sharat Chandran	Yang He	Kevin Liu	Ameya Prabhu	Shubham Tulsiani	Linguang Zhang
XiaoJun Chang	Wolfgang Heidrich	Tsung-Yu Lin	True Price	Dimitrios Tzionas	Qiangqong Zhang
Binghui Chen	Lisa Anne Hendricks	Roei Litman	Véronique Prinet	Phani Krishna Uppala	Shanshan Zhang
Dongdong Chen	Steven Hickson	Buyu Liu	Senthil Purushwalkam	Tuan-Hung VU	Yongqiang Zhang
Kan Chen	Tsung-Ying Ho	Liu liu	Yuankai Qi	Chuan Wang	Zheng Zhang
Shixing Chen	Namdar Homayounfar	Mengyuan Liu	Siyuan Qiao	Dequan Wang	Zhijun Zhang
Xinlei Chen	Guosheng Hu	Sheng Liu	Jie Qin	Di Wang	Zhishuai Zhang
Yi-Ting Chen	Peiyun Hu	Shu Liu	Venkatesh Babu Radhakrishnan	Guangrun Wang	Kai Zhao
Gong Cheng	Binh-Son Hua	Siqi Liu	Ilija Radosavovic	Hongxing Wang	Qijun Zhao
Jun Cheng	Lei Huang	Weifeng Liu	Yongming Rao	Jian Wang	Xiangyun Zhao
Wei-Chen Chiu	Qingqiu Huang	Weiyang Liu	Zhile Ren	Kang Wang	Xin Zhao
Donghyeon Cho	Weilin Huang	Xiaofeng Liu	Christian Richardt	Lei Wang	Zhun Zhong
Jin Young Choi	Yue Huang	Yaojie Liu	Hayko Riemenschneider	Pei Wang	Hao Zhou
Hang Chu	Junhwa Hur	Yebin Liu	Mikel Rodriguez	Wenlin Wang	Ning Zhou
Canton Cristian	Nazli Ikizler-Cinbis	Yun Liu	Michal Rolinek	Xiang Wang	Xingyi Zhou
Yin Cui	Vamsi Ithapu	Zhijian Liu	Xuejian Rong	Xiaosong Wang	Ji Zhu
Yuchao Dai	Suyog Jain	Roberto Lopez-Sastre	Amir Rosenfeld	Xin Wang	Xizhou Zhu
Zachary Daniels	Shihao Ji	Hongtao Lu	Peter Roth	Xinggang Wang	Yixin Zhu
Fillipe D M de Souza	Kui Jia	Ruotian Luo	Christian Rupprecht	Yangang Wang	Zheng Zhu
Koichiro Deguchi	Xudong Jiang	Chih-Yao Ma	Hideo Saito	Zhe Wang	Bingbing Zhuang
Luca Del Pero	Jianbo Jiao	Lin Ma	Fatemeh Sadat Saleh	Ping Wei	Chuhang Zou
Ilke Demir	SouYoung Jin	Michael Maire	Aswini Sankaranarayanan	Michael Weinmann	
Jiankang Deng	Ole Johannsen	Clement Mallet	Swami Sankaranarayanan	Bihan Wen	
Frédéric Devernavy	Justin Johnson	Massimiliano Mancini	Torsten Sattler	Williem Williemi	
Ferran Diego	Kushal Kafle	David Masip	Alexander Sax	Christian Wolf	
Bo Dong	Zdenek Kalal	Roey Mechrez	Vivek Sharma	Sanghyun Woo	

Saturday, June 15

1700–2000 **Registration** (Promenade Atrium & Plaza)

Sunday, June 16

NOTE: Use the QR code for each tutorial's website for more information on that tutorial. Here's the QR code to the CVPR Tutorials page.



0700–1730 **Registration** (Promenade Atrium & Plaza)

0730–0900 **Breakfast** (Pacific Ballroom)

1000–1045 **Morning Break** (Exhibit Hall)

1200–1330 **Lunch** (Pacific Ballroom)

1515–1600 **Afternoon Break** (Exhibit Hall)

Tutorial: Towards Relightable Volumetric Performance Capture of Humans

Organizers: Sean Fanello Sofien Bouaziz
 Christoph Rhemann Paul Debevec
 Graham Fyffe Shahram Izadi
 Jonathan Taylor

Location: 201A
Time: Full Day (0900-1730)

Description: Volumetric (4D) performance capture is fundamental for AR/VR content generation. Designing a volumetric capture pipeline involves developing high quality sensors and efficient algorithms that can leverage new and existing sensing technology. To this end, we leverage a combination of active sensors with traditional photometric stereo methods. As a result, we have developed a wide range of high quality algorithms for reconstruction, tracking and texturing of humans in 4D.

In this tutorial we will walk the attendee through the ins and outs of building such a system from the ground up. In the first part of this we will consider the hardware design choices for cameras, sensors, lighting, and depth estimation algorithms. We then walk through the proposed RGBD active sensors to achieve high quality results with reasonable runtime. In the second part we will cover reconstruction and tracking techniques for people. We will review state of the art algorithms such as Kinect Fusion, Dynamic Fusion, Fusion4D and Motion2Fusion. We will also detail parametric tracking approaches for faces, hands and bodies. In the third part we will focus on photometric stereo, texturing and relightability: we will detail the state of the art algorithm together with our choices. Finally we will discuss some applications and capabilities that 3D capture technologies enable. We will put emphasis on virtual and augmented reality scenarios and highlight recent trends in machine learning that aim at replacing traditional graphics pipelines.



Tutorial: Deep Learning for Content Creation

Organizers: Deqing Sun
 Ming-Yu Liu
 Orazio Gallo
 Jan Kautz

Location: 204
Time: Full Day (0845-1700)

Description: Content creation has several important applications ranging from virtual reality, videography, gaming, and even retail and advertising. The recent progress of deep learning and machine learning techniques allowed to turn hours of manual, painstaking content creation work into minutes or seconds of automated work. For instance, generative adversarial networks (GANs) have been used to produce photorealistic images of items such as shoes, bags, and other articles of clothing, interior/industrial design, and even computer games' scenes. Neural networks can create impressive and accurate slow-motion sequences from videos captured at standard frame rates, thus side-stepping the need for specialized and expensive hardware. Style transfer algorithms can convincingly render the content of one image with the style of another and offer unique opportunities for generating additional and more diverse training data—in addition to creating awe-inspiring, artistic images. Learned priors can also be combined with explicit geometric constraints, allowing for realistic and visually pleasing solutions to traditional problems such as novel view synthesis, in particular for the more complex cases of view extrapolation.

Deep learning for content creation lies at the intersection of graphics and computer vision. However, researchers and professionals in either field may not be aware of its full potential and inner workings. This tutorial has several goals. First, it will cover some introductory concepts to help interested researchers from other fields get started in this exciting new area. Second, it will present selected success cases to advertise how deep learning can be used for content creation. More broadly, it will serve as a forum to discuss the latest topics in content creation and the challenges that vision and learning researchers can help solve. With its many confirmed speakers, the tutorial will comprise three parts: theoretical foundations, image synthesis, and video synthesis.



Tutorial: Visual Recognition and Beyond

Organizers: Christoph Feichenhofer Georgia Gkioxari
 Kaiming He Alexander Kirillov
 Ross Girshick Piotr Dollar

Location: 104C
Time: Half Day - Morning (0830-1200)

Description: This tutorial covers topics at the frontier of research on visual recognition. We will discuss the recent advances on instance-level recognition from images and videos, covering in detail the most recent work in the family of visual recognition tasks. The talks cover image classification, video classification, object detection, action detection, instance segmentation, semantic segmentation, panoptic segmentation, and pose estimation.



Tutorial: Vision Meets Mapping: Computer Vision for Location-Based Reasoning and Mapping

Organizers: Xiang Ma

Location: Seaside 6

Time: Half Day - Afternoon (1300-1730)

Description: We live in a world where location is the key for many activities, and computer vision plays a key role for answering the fundamental questions for location-based reasoning and mapping: (1) Where am I? (2) Where i am going? (3) How do i get there? The future world, not too far from our real life, would be full of autonomous robotics and AI agents: autonomous vehicles, autonomous taxis, autonomous delivery vans, autonomous drones etc. and location-based reasoning, especially mapping, has drawn a lot of attention from the computer vision community recently, both academia and industry.

This tutorial session brings together people from around the world who are practicing computer vision research for mapping/location-based reasoning in both industry and academia, and would cover the following topics of interests (including but not limited to): vision-based map making, vision-based high definition map creation, crowd-sourced map creation, semantic map, structure map, vision-based localization, lidar-based localization, multi-sensor-based localization, 2D/3D scene understanding for location-based reasoning and 2D/3D visual landmark detection.



Tutorial: Action Classification and Video Modelling

Organizers: Efstratios Gavves

Joao Carreira

Christoph Feichtenhofer

Lorenzo Torresani

Basura Fernando

Location: 202B

Time: Half Day - Afternoon (1300-1700)

Description: Deep Learning has been a great influential in most Computer Vision and Machine Learning tasks. In video analysis problems, such as action recognition and detection, motion analysis and tracking, the progress has arguably been slower, with shallow models remaining surprisingly competitive. Recent developments have demonstrated that careful model design and end-to-end model training, as well as large and well-annotated datasets, have finally led to strong results using deep architectures for video analysis. However, the details and the secrets to achieve good accuracies with deep models are not always transparent. Furthermore, it is not always clear whether the networks resulting from end-to-end training are truly providing better video models or if instead are simply overfitting their large capacity to the idiosyncrasies of each dataset.

In recent years there have been many conference workshops on action classification and video recognition, but very few tutorials on this topic. This has been likely due to the lack of a common successful methodology toward video modeling that could be taught coherently in a tutorial setting. However, recent advances in deep learning for video and the emergence of established winning recipes in this area warrant a tutorial on video modeling and action classification.



Tutorial: Capsule Networks for Computer Vision

Organizers: Yogesh Singh Rawat

Mubarak Shah

Ulas Bagci

Sara Sabour

Rodney LaLonde

Kevin Duarte

Location: 104C

Time: Half Day - Afternoon (1300-1700)

Description: A capsule network provides an effective way to model part-to-whole relationships between entities and allows to learn viewpoint invariant representations. Through this improved representation learning, capsule networks are able to achieve good performance in multiple domains with a drastic decrease in the number of parameters. Recently, capsule networks have shown state-of-the-art results for human action localization in a video, object segmentation in medical images, and text classification. This tutorial will provide a basic understanding of capsule network, and we will discuss its use in a variety of computer vision tasks such as image classification, object segmentation, and activity detection.



Tutorial: Distributed Private Machine Learning for Computer Vision: Federated Learning, Split Learning and Beyond

Organizers: Brendan McMahan

Ramesh Raskar

Jakub Konečný

Otkrist Gupta

Hassan Takabi

Praneeth Vepakomma

Location: 203A

Time: Half Day - Afternoon (1330-1730)

Description: This tutorial presents different methods for protecting confidential data on clients while still allowing servers to train models. In particular, we focus on distributed deep learning approaches under the constraint that local data sources of clients (e.g. photos on phones or medical images at hospitals) are not allowed to be shared with the server or amongst other clients due to privacy, regulations or trust. We describe such methods that include federated learning, split learning, homomorphic encryption, and differential privacy for securely learning and inferring with neural networks. We also study their trade-offs with regards to computational resources and communication efficiency in addition to sharing practical know-how of deploying such systems. We discuss practical software solutions available to computer vision researchers.



Tutorial: Perception at Magic Leap

Organizers: Ashwin Swaminathan

Jean-Yves Bouguet

Dan Farmer

David Molyneaux

Location: 201B

Time: Half Day - Afternoon (1300-1700)

Description: This tutorial presents the importance of Computer Vision and Deep learning techniques in making Magic Leap an effective spatial computing platform. The four fundamental modalities are introduced: head pose tracking, world reconstruction, eye tracking and hand tracking; emphasizing on the two main general themes: Understanding the world (spatial localization, environment mapping) and Understanding user's intent (eye, gaze and hands). The tutorial will provide a deep dive into the main modalities along with key challenges and open problems.



Tutorial: Recent Advances in Visual Data Summarization

Organizers: Rameswar Panda
Ehsan Elhamifar
Amin Karbasi
Michael Gygli



Location: 203C

Time: Half Day - Afternoon (1330-1730)

Description: Visual data summarization has many applications ranging from computer vision (video summarization, video captioning, active visual learning, object detection, image/video segmentation, etc) to data mining (recommender systems, web-data analysis, etc). As a consequence, new important research topics and problems are recently appearing, (i) online and distributed summarization, (ii) summarization with privacy and fairness constraints, (iii) weakly supervised summarization, (iv) summarization in sequential data, as well as (v) summarization in networks of cameras, in particular, for surveillance tasks. The objective of this tutorial is to present the audience with a unifying perspective of the visual data summarization problem from both theoretical and application standpoint, as well as to discuss, motivate and encourage future research that will spur disruptive progress in the the emerging field of summarization.

Notes:

Grid area for taking notes on the left side of the page.

Large grid area for taking notes on the right side of the page.

Sunday, June 16

NOTE: Use the QR code for each workshop's website to find the workshop's schedule. Here's the QR code to the CVPR Workshops page.



0700-1730 Registration (Promenade Atrium & Plaza)

0730-0900 Breakfast (Pacific Ballroom)

1000-1045 Morning Break (Exhibit Hall; Hyatt Foyer)

1200-1330 Lunch (Pacific Ballroom)

1515-1600 Afternoon Break (Exhibit Hall; Hyatt Foyer)

3D Scene Generation

Organizers: Daniel Ritchie
Angel X. Chang
Qixing Huang
Manolis Savva



Location: 103A
Time: 0845-1740 (Full Day)

Summary: People spend a large percentage of their lives indoors--in bedrooms, living rooms, offices, kitchens, etc.--and the demand for virtual versions of these real-world spaces has never been higher. Game developers, VR/AR designers, architects, and interior design firms all increasingly use virtual 3D scenes for prototyping and final products. Furthermore, AI/vision/robotics researchers are also turning to virtual environments to train data-hungry models for visual navigation, 3D reconstruction, activity recognition, and more.

As the vision community turns from passive internet-images-based vision tasks to applications such as the ones listed above, the need for virtual 3D environments becomes critical. The community has recently benefited from large scale datasets of both synthetic 3D environments and reconstructions of real spaces, as well as 3D simulation frameworks for studying embodied agents. While these existing datasets are valuable, they are also finite in size and don't adapt to the needs of different vision tasks. To enable large-scale embodied visual learning in 3D environments, we must go beyond static datasets and instead pursue the automatic synthesis of novel, task-relevant virtual environments.

In this workshop, we aim to bring together researchers working on automatic generation of 3D environments for computer vision research with researchers who use 3D environment data for computer vision tasks. We define "generation of 3D environments" to include methods that generate 3D scenes from sensory inputs (e.g. images) or from high-level specifications (e.g. "a chic apartment for two people"). Vision tasks that consume such data include automatic scene classification and segmentation, 3D reconstruction, human activity recognition, robotic visual navigation, and more.

Language and Vision

Organizers: Siddharth Narayanaswamy
Andrei Barbu
Dan Gutfreund
Philip Torr



Location: Seaside 1
Time: TBA (Full Day)

Summary: The interaction between language and vision, despite seeing traction as of late, is still largely unexplored. This is a particularly relevant topic to the vision community because humans routinely perform tasks which involve both modalities. We do so largely without even noticing. Every time you ask for an object, ask someone to imagine a scene, or describe what you're seeing, you're performing a task which bridges a linguistic and a visual representation. The importance of vision-language interaction can also be seen by the numerous approaches that often cross domains, such as the popularity of image grammars. More concretely, we've recently seen a renewed interest in one-shot learning for object and event models. Humans go further than this using our linguistic abilities; we perform zero-shot learning without seeing a single example. You can recognize a picture of a zebra after hearing the description "horse-like animal with black and white stripes" without ever having seen one.

Furthermore, integrating language with vision brings with it the possibility of expanding the horizons and tasks of the vision community. A major difference between human and machine vision is that humans form a coherent and global understanding of a scene. This process is facilitated by our ability to affect our perception with high-level knowledge which provides resilience in the face of errors from low-level perception. It also provides a framework through which one can learn about the world: language can be used to describe many phenomena succinctly thereby helping filter out irrelevant details.

Vision for All Seasons: Bad Weather and Nighttime

Organizers: Dengxin Dai
Christos Sakaridis
Robby T. Tan
Radu Timofte
Daniel Olmeda Reino
Wim Abbeloos
Nicolas Vignard
Marc Proesmans
Jiri Matas
Roberto Cipolla
Bernt Schiele
Luc Van Gool

Location: 102B
Time: 0900-1730 (Full Day)



Summary: Adverse weather and illumination conditions (e.g. fog, rain, snow, ice, low light, nighttime, glare and shadows) create visibility problems for the sensors that power automated systems. Many outdoor applications such as autonomous cars and surveillance systems are required to operate smoothly in the frequent scenarios of bad weather. While rapid progress is being made in this direction, the performance of current vision algorithms is still mainly benchmarked under clear weather conditions (good weather, favorable lighting). Even the top-performing state-of-the-art algorithms undergo a severe performance degradation under adverse conditions. The aim of this workshop is to bring together bright minds to share knowledge and promote research into the design of robust vision algorithms for adverse weather and illumination conditions.

Deep-Vision: New Frontiers and Advances in Theory in Deep Learning for Computer Vision

Organizers: Jose M. Alvarez
Cristian Canton
Yann LeCun

Location: Terrace Theater

Time: 0900-1800 (Full Day)



Summary: Deep Learning has become the standard tool to approach almost any computer vision problem. In order to present an attractive workshop that brings a different perspective to the topic, we are focusing this edition on brave new ideas, theoretical discussions and, in general, imaginative approaches that may yield to the next family of deep learning algorithms.

We encourage researchers to formulate innovative learning theories, feature representations, and end-to-end vision systems based on deep learning. We also encourage new theories and processes for dealing with large scale image datasets through deep learning architectures.

Computer Vision for Global Challenges

Organizers: Laura Sevilla-Lara Mourad Gridach
Yannis Kalantidis Anna Lerner
Maria De-Arteaga Amir Zamir
Timnit Gebru Stefano Ermon
Kris Sankaran Lorenzo Torresani
Ernest Mwebaze Larry Zitnick
John Quinn Jitendra Malik
Mutembesa Daniel

Location: Grand Ballroom A

Time: 0845-1800 (Full Day)



Summary: Computer vision and its applications are often strongly tied. While many of our tasks and datasets are designed to be fundamental, others take a specific application as problem motivation or source of data. With much of the research and applications concentrated in a handful of advanced markets, biases are likely to emerge in the datasets, tasks and ultimately the direction of the advancement of the field.

Widening the scope of computer vision applications to address global challenges could lead to a win for the computer vision community, the organizations and individuals working to address global challenges, and the world at large: vision could positively impact the lives of 6 billion people in emerging markets, e.g. by developing applications for agriculture, digital health care delivery, or disaster readiness - and these populations could help reveal some of the blind spots and biases in the current computer vision datasets, tasks, and practices.

In this workshop we propose to take a first step towards bridging the gap between computer vision and global challenges by:

- Finding intellectually interesting challenges that broaden the scope of computer vision problems through new datasets and tasks related to development priorities like the UN’s Sustainable Development Goals
- Identifying computer vision techniques that can help solve problems with large positive societal impact in emerging markets
- To give individuals, universities and organizations the ability to contribute to the previous two goals, through collaborations, mentorship, research grants, etc.

The Bright and Dark Sides of Computer Vision: Challenges and Opportunities for Privacy and Security

Organizers: David Crandall
Jan-Michael Frahm
Mario Fritz
Apu Kapadia
Vitaly Shmatikov

Location: 102C

Time: 0900-1800 (Full Day)



Summary: Computer vision is finally working in the real world, but what are the consequences on privacy and security? For example, recent work shows that vision algorithms can spy on smartphone keypresses from meters away, steal information from inside homes via hacked cameras, exploit social media to de-anonymize blurred faces, and reconstruct images from features like SIFT. Vision could also enhance privacy and security, for example through assistive devices for people with disabilities, phishing detection techniques that incorporate visual features, and image forensic tools. Some technologies present both challenges and opportunities: biometrics techniques could enhance security but may be spoofed, while surveillance systems enhance safety but create potential for abuse.

We need to understand the potential threats and opportunities of vision to avoid creating detrimental societal effects and/or facing public backlash. Following up on last year's very successful workshops at CVPR 2017, and CVPR 2018, this workshop will continue to explore the intersection between computer vision and security and privacy to address these issues.

BioImage Computing

Organizers: Dagmar Kainmueller
Kristin Branson
Jan Funke
Florian Jug
Anna Kreshuk
Carsten Rother

Location: Seaside 3

Time: 0900-1730 (Full Day)



Summary: Bio-image computing (BIC) is a rapidly growing field at the interface of engineering, biology and computer science. Advanced light microscopy can deliver 2D and 3D image sequences of living cells with unprecedented image quality and ever increasing resolution in space and time. The emergence of novel and diverse microscopy modalities has provided biologists with unprecedented means to explore cellular mechanisms, embryogenesis, and neural development, to mention only a few fundamental biological questions. Electron microscopy provides information on the cellular structure at nanometer resolution. Here, correlating light microscopy and electron microscopy at the subcellular level, and relating both to animal behavior at the macroscopic level, is of paramount importance. The enormous size and complexity of these data sets, which can exceed multiple TB per volume or video, requires state-of-the-art computer vision methods.

This workshop brings the latest challenges in bio-image computing to the computer vision community. It will showcase the specificities of bio-image computing and its current achievements, including issues related to image modeling, denoising, super-resolution, multi-scale instance- and semantic segmentation, motion estimation, image registration, tracking, classification, event detection -- important topics that appertain to the computer vision field.

Adversarial Machine Learning in Real-World Computer Vision Systems

Organizers: Bo Li
Li Erran Li
David Forsyth
Dawn Song
Ramin Zabih
Chaowei Xiao



Location: Seaside 5
Time: 0800-1615 (Full Day)

Summary: As computer vision models are being increasingly deployed in the real world, including applications that require safety considerations such as self-driving cars, it is imperative that these models are robust and secure even when subject to adversarial inputs. This workshop will focus on recent research and future directions for security problems in real-world machine learning and computer vision systems. We aim to bring together experts from the computer vision, security, and robust learning communities in an attempt to highlight recent work in this area as well as to clarify the foundations of secure machine learning. We seek to come to a consensus on a rigorously framework to formulate adversarial machine learning problems in computer vision, characterize the properties that ensure the security of perceptual models, and evaluate the consequences under various adversarial models. Finally, we hope to chart out important directions for future work and cross-community collaborations, including computer vision, machine learning, security, and multimedia communities.

Medical Computer Vision

Organizers: Tal Arbel
Le Lu
Leo Grady
Nicolas Padov



Location: Seaside 4
Time: 0830-1745 (Full Day)

Summary: Machine learning methods have become very popular in recent MICCAI conferences. Deep learning (auto-encoder, convolutional neural networks, deep reinforcement learning) has revolutionized the field of computer vision, but faces particular challenges in the field of medical image analysis. In addition to the shortage of large, annotated datasets, other challenges are presented as medical images are sparse, real "ground truth" labels are usually non-existent, segmentation of very small structures are required, etc. Development of new computer vision representations and frameworks are needed to address the particular needs of the medical communities. This workshop will cover all topics related to the development of novel machine learning in medical imaging, particularly for "big clinical data". The workshop will consist of a series of 30 minute invited talks from researchers working in both fields. We have lined up an excellent list of speakers, from academia and from industry, who will present work ranging from the more theoretical to the more applied. This workshop aims to bridge the gap between the medical image analysis/computer aided intervention communities and the computer vision communities, providing a forum for exchanges of ideas and potential new collaborative efforts, encourage more data sharing, radiology image database building and information exchange on machine learning systems for medical applications. This collective effort among peers will facilitate the next level of large scale machine learning methods.

Embedded Vision (Special Theme: UAVs)

Organizers: Martin Humenberger
Tse-Wei Chen
Rajesh Narasimha
Stephan Weiss
Roland Brockers



Location: 101A
Time: 0900-1730 (Full Day)

Summary: UAVs with embedded and real-time vision processing capabilities have proven to be highly versatile autonomous robotic platforms in a variety of environments and for different tasks. The workshop aims at collecting and discussing next-generation approaches for online and onboard visual systems embedded on UAVs in order to elaborate a holistic picture of the state of the art in this field, and also to reveal remaining challenges and issues. The discussion of these topics will identify next steps in this research area for both young researchers as well as senior scientists. The invited speakers from academia will highlight past and current research and the invited speakers from industry will help closing the loop how research efforts are and will be perceived for industrial applications. Topics of interest include vision-based navigation, vision-based (multi-)sensor-fusion, online reconstruction, collaborative vision algorithms for navigation and environment perception, and related topics.

Visual Understanding by Learning From Web Data

Organizers: Wen Li
Limin Wang
Wei Li
Eirikur Agustsson

Jesse Berent
Abhinav Gupta
Rahul Sukthankar
Luc Van Gool

Location: 203B
Time: 0830-1600 (Full Day)

Summary: The recent success of deep learning has shown that a deep architecture in conjunction with abundant quantities of labeled training data is the most promising approach for many vision tasks. However, annotating a large-scale dataset for training such deep neural networks is costly and time-consuming, even with the availability of scalable crowdsourcing platforms like Amazon's Mechanical Turk. As a result, there are relatively few public large-scale datasets (e.g., ImageNet and Places2) from which it is possible to learn generic visual representations from scratch.

It is unsurprising that there is continued interest in developing novel deep learning systems that trained on low-cost data for image and video recognition tasks. Among different solutions, crawling data from Internet and using the web as a source of supervision for learning deep representations has shown promising performance for a variety of important computer vision applications. However, the datasets and tasks differ in various ways, which makes it difficult to fairly evaluate different solutions, and identify the key issues when learning from web data.

This workshop promotes the advance of learning state-of-the-art visual models directly from the web, and bringing together computer vision researchers in this field. We release a large scale web image dataset named WebVision. The datasets consists of 16 million of web images crawled from Internet for 5,000 visual concepts. A validation set consists of around 290K images with human annotation will be provided for the convenience of algorithmic development.



Semantic Information

Organizers: René Vidal
John Shawe-Taylor

Location: 202C

Time: 0845-1700 (Full Day)

Summary: Classical notions of information, such as Shannon entropy, measure the amount of information in a signal in terms of the frequency of occurrence of symbols. Such notions are very useful for tasks such as data compression. However, they are not as useful for tasks such as visual recognition, where the semantic content of the scene is essential. Moreover, classical notions of information are affected by nuisance factors, such as viewpoint and illumination conditions, which are irrelevant to the recognition task. The goal of this workshop is to bring together researchers in computer vision, machine learning and information theory, to discuss recent progress on defining and computing new notions of information that capture the semantic content of multi-modal data. Topics of interest include but are not limited to information theoretic approaches to scene understanding, representation learning, domain adaptation, and generative adversarial networks as well as the interplay between information and semantic content.



Analysis and Modeling of Faces and Gestures

Organizers: Joseph P. Robinson Sarah Ostadabbas
Yun (Raymond) Fu Sheng Li
Ming Shao Zhengming Ding
Siyu Xia

Location: 202A

Time: 0830-1700 (Full Day)

Summary: We have experienced rapid advances in face, gesture, and cross-modality (e.g., voice and face) technologies. This is due to the deep learning and large-scale, labeled image collections. The progress made in deep learning continues to push renowned public databases to near saturation which, thus, calls more evermore challenging image collections to be compiled as databases. In practice, and even widely in applied research, off-the-shelf deep learning models have become the norm, as numerous pre-trained networks are available for download and are readily deployed to new, unseen data. We have almost grown "spoiled" from such luxury, which, in all actuality, has enabled us to stay hidden from many truths. Theoretically, the truth behind what makes neural networks more discriminant than ever before is still, in all fairness, unclear—rather, they act as a sort of black box to most practitioners and even researchers, alike. More troublesome is the absence of tools to quantitatively and qualitatively characterize existing deep models. With the frontier moving forward at rates incomparable to any spurt of the past, challenges such as high variations in illuminations, pose, age, etc., now confront us. However, state-of-the-art deep learning models often fail when faced with such challenges owed to the difficulties in modeling structured data and visual dynamics. This workshop provides a forum for researchers to review the recent progress of recognition, analysis, and modeling of face, body, and gesture, while embracing the most advanced deep learning systems available for face and gesture analysis, particularly, under an unconstrained environment like social media and across modalities like face to voice.



Augmented Human: Human-Centric Understanding and 2D/3D Synthesis

Organizers: Xiaodan Liang
Haoye Dong
Yunchao Wei
Xiaohui Shen
Jiashi Feng
Song-Chun Zhu

Location: 102A

Time: 0830-1700 (Full Day)

Summary: An ultimate goal of computer vision is to augment humans in a variety of application fields. Developing solutions to comprehensive human-centric visual applications in the wild scenarios, regarded as one of the most fundamental problems in computer vision, could have a crucial impact in many industrial application domains, such as virtual reality, human-computer interaction, human motion analysis, and advanced robotic perception. Human-centric understanding, including human parsing/detection, pose estimation, and relationship detection, are often regarded as the very first step for higher-level activity/event recognition and detection. Nonetheless, a large gap seems to exist between what is needed by real-life applications and what is achievable from modern computer vision techniques. Further, more virtual reality and 3D graphic analysis research advances are urgently expected for advanced human-centric analysis. For example, virtual 2D/3D clothes "try-on" systems that seamlessly fits various clothes into 3D human body shape has attracted several commercial interests. Human motion synthesis and prediction can bridge the virtual and real worlds, such as, simulating virtual characters to mimic the human behaviors or empowering more intelligent robotic interactions with humans by enabling causal inferences for human activities. The goal of this workshop is to allow researchers from the fields of human-centric understanding and 2D/3D synthesis to present their progress, communication and co-develop novel ideas that potentially shape the future of this area and further advance the performance and applicability of correspondingly built systems in real-world conditions.



3D HUMANS: HUMAN pose Motion Activities aND Shape in 3D

Organizers: Grégory Rogez
Javier Romero
Manuel J. Marin-Jiménez

Location: 104B

Time: 0850-1700 (Full Day)

Summary: This workshop aims at gathering researchers who work on 3D understanding of humans from visual data, including topics such as 3D human pose estimation and tracking, 3D human shape estimation from RGB images or human activity recognition from 3D skeletal data. Current computer vision algorithms and deep learning-based methods can detect people in images and estimate their 2D pose with a remarkable accuracy. However, understanding humans and estimating their pose and shape in 3D is still an open problem. The ambiguities in lifting 2D pose to 3D, the lack of annotated data to train 3D pose regressors in the wild and the absence of a reliable evaluation dataset in real world situations make the problem very challenging. The workshop will include several high quality invited talks and a poster session with invited posters.



Learning From Unlabeled Videos

Organizers: Yale Song
 Carl Vondrick
 Katerina Fragkiadaki
 Honglak Lee
 Rahul Sukthankar

Location: Hyatt Regency H
Time: 0845-1705 (Full Day)



Summary: Deep neural networks trained with a large number of labeled images have recently led to breakthroughs in computer vision. However, we have yet to see a similar level of breakthrough in the video domain. Why is this? Should we invest more into supervised learning or do we need a different learning paradigm?

Unlike images, videos contain extra dimensions of information such as motion and sound. Recent approaches leverage such signals to tackle various challenging tasks in an unsupervised/self-supervised setting, e.g., learning to predict certain representations of the future time steps in a video-- (RGB frame, semantic segmentation map, optical flow, camera motion, and corresponding sound), learning spatio-temporal progression from image sequences, and learning audio-visual correspondences.

This workshop aims to promote comprehensive discussion around this emerging topic. We invite researchers to share their experiences and knowledge in learning from unlabeled videos, and to brainstorm brave new ideas that will potentially generate the next breakthrough in computer vision.

Autonomous Driving – Beyond Single-Frame Perception

Organizers: Li Erran Li
 Marc Pollefeys
 Daniela Rus
 Kilian Weiberger
 Ruigang Yang

Location: Grand Ballroom B
Time: 0900-1730 (Full Day)



Summary: The CVPR 2019 Workshop on Autonomous Driving — beyond single frame perception builds on 2018 with a focus on multi-frame perception, prediction, and planning for autonomous driving. It aims to bring together researchers and engineers from academia and industry to discuss computer vision applications in autonomous driving. In this one day workshop, we will have invited speakers, panel discussions, and technical benchmark challenges to present the current state of the art, as well as the limitations and future directions for computer vision in autonomous driving, arguably the most promising application of computer vision and AI in general.

Vision Meets Cognition

Organizers: Jiajun Wu Zhoutong Zhang
 Yixin Zhu Chenfanfu Jiang
 Yunzhi Li Song-Chun Zhu
 Siyuan Qi Joshua B. Tenenbaum

Location: 103C
Time: 0900-1730 (Full Day)



Summary: The Vision Meets Cognition (VMC) workshop will bring together researchers from computer vision, graphics, robotics, cognitive science, and developmental psychology to advance computer vision systems toward cognitive understanding of visual data.

Bridging the Gap Between Computational Photography and Visual Recognition

Organizers: Walter J. Scheirer Kevin Bowyer
 Zhangyang (Atlas) Wang Thomas S. Huang
 Jiaying Liu Rosaura VidalMata
 Wenqi Ren Sreya Banerjee
 Wenhan Yang Ye Yuan

Location: Hyatt Regency C
Time: 0830-1800 (Full Day)



Summary: The advantages of collecting images from outdoor camera platforms, like UAVs, surveillance cameras and outdoor robots, are evident and clear. For instance, man-portable UAV systems can be launched from safe positions to survey difficult or dangerous terrain, acquiring hours of video without putting human lives at risk. What is unclear is how to automate the interpretation of these images - a necessary measure in the face of millions of frames containing artifacts unique to the operation of the sensor and optics platform in outdoor, unconstrained, and usually visually degraded environments. Continuing the success of the 1st UG² Prize Challenge workshop held at CVPR 2018, UG²+ provides an integrated forum for researchers to review the recent progress of handling various adverse visual conditions in real-world scenes, in robust, effective and task-oriented ways. Beyond the human vision-driven restorations, we also extend particular attention to the degradation models and the related inverse recovery processes that may benefit successive machine vision tasks. We embrace the most advanced deep learning systems, but are still open to classical physically grounded models, as well as any well-motivated combination of the two streams.

AI City Challenge

Organizers: Milind Naphade Xiaodong Yang
 Anuj Sharma Sywei Lyu
 David Anastasiu Rama Chellappa
 Ming-Yu Liu Jenq-Neng Hwang
 Ming-Ching Chang

Location: 101B
Time: 0900-1730 (Full Day)



Summary: Immense opportunity exists to make transportation systems smarter, based on sensor data from traffic, signaling systems, infrastructure, and transit. Unfortunately, progress has been limited for several reasons — among them, poor data quality, missing data labels, and the lack of high-quality models that can convert the data into actionable insights. There is also a need for platforms that can handle analysis from the edge to the cloud, which will accelerate the development and deployment of these models.

The AI City Challenge Workshop at CVPR 2019 addresses these challenges by encouraging research and development into techniques that rely less on supervised approaches and more on transfer learning, unsupervised and semi-supervised approaches that go beyond bounding boxes. It will focus on Intelligent Transportation System (ITS) problems, such as:

- City-scale multi-camera vehicle tracking
- City-scale multi-camera vehicle re-identification
- Traffic anomaly detection – Leveraging unsupervised learning to detect anomalies such as lane violation, illegal U-turns, wrong-direction driving, etc.

CARLA Autonomous Driving Challenge

Organizers: German Ros
 Vladlen Koltun
 Alexey Dosovitskiy
 David Vázquez
 Felipe Codevilla
 Antonio M. López



Location: Seaside Ballroom A
Time: 0900-1700 (Full Day)

Summary: The field of mobility is going through a revolution that is changing how we understand transportation. The potential benefits of autonomous vehicles are immense: elimination of accidents caused by human errors, a reduction in carbon dioxide emission, more efficient use of energy and infrastructure, among others. However, many technical questions remain unanswered. What is the best path towards fully autonomous driving? What is the optimal combination of sensors for driving? Which components of the driving stack should be hand-crafted and which can be learned from data? In which traffic situations do different algorithms fail?

Despite the tremendous interest in autonomous driving, there are no clear answers to these questions. The reason is that evaluation of driving systems in the real world is extremely costly and thus only available to large corporations. Thus, until recently there was no open and accessible way to evaluate autonomous driving algorithms. To counter this situation and expedite research in the field of autonomous driving, we present the CARLA Autonomous Driving Challenge, a step towards democratizing autonomous driving research and development. Participants will deploy state-of-the-art autonomous driving systems to tackle complex traffic scenarios in CARLA — an open source driving simulator. CARLA provides an even playing field for all participants: every vehicle will face the same set of traffic situations and challenges. This allows for a fair and reliable comparison of various autonomous driving approaches.

Weakly Supervised Learning for Real-World Computer Vision Applications and Learning From Imperfect Data Challenge

Organizers: Yunchao Wei
 Shuai (Kyle) Zheng
 Ming-Ming Cheng
 Xiaodan Liang
 Hang Zhao

Honghui Shi
 Liwei Wang
 Antonio Torralba
 Thomas Huang



Location: Hyatt Regency B
Time: 0820-1620 (Full Day)

Summary: Weakly supervised learning attempts to address the challenging pattern recognition tasks by learning from weak or imperfect supervision. Supervised learning methods including Deep Convolutional Neural Networks (DCNNs) have significantly improved the performance in many problems in the field of computer vision, thanks to the rise of large-scale annotated data sets and the advance in computing hardware. However, these supervised learning approaches are notorious data hungry, which sometimes makes them impractical in many real-world industrial applications. We often face the problem that we can't acquire sufficient perfect annotations for reliable training models. To address this problem, many efforts in weakly supervised learning approaches have been made to improve the DCNNs training to deviate from traditional paths of supervised learning using imperfect data. For instance,

various approaches have proposed new loss functions or novel training schemes. Weakly supervised learning is a popular research direction in computer vision and machine learning communities. This workshop investigates current methods of building industry level AI system relying on learning from imperfect data. We hope this workshop will attract attention and discussions from both industry and academic people.

Habitat: Embodied Agents Challenge

Organizers: Manolis Savva
 Abhishek Kadian
 Oleksandr Maksymets
 Erik Wijmans
 Bhavana Jain
 Julian Straub
 Abhishek Das
 Samyak Datta

Georgia Gkioxari
 Marcus Rohrbach
 Stefan Lee
 Peter Anderson
 Vladlen Koltun
 Jitendra Malik
 Devi Parikh
 Dhruv Batra

Location: Seaside Ballroom B
Time: 0900-1810 (Full Day)

Summary: There has been a recent shift in the computer vision community from tasks focusing on internet images to active settings involving embodied agents that perceive and act within 3D environments. Practical deployment of AI agents in the real world requires study of active perception and coupling of perception with control as in embodied agents.



This workshop has two objectives. First, to establish a unified platform and a set of benchmarks to measure progress in embodied agents by hosting an embodied agents challenge named Habitat Challenge. These benchmarks will evaluate algorithms for navigation and question-answering tasks using the Habitat platform. The benchmarks and unified embodied agents platform will catalyze future work, promoting reproducibility, reusability of code, and consistency in evaluation.

The second objective of the workshop is to bring together researchers from the fields of computer vision, language, graphics, and robotics to share work being done in the area of multi-modal AI, as well as discuss future directions in research on embodied agents connecting several recent threads of research on: visual navigation, natural language instruction following, embodied question answering and language grounding.

Deep Learning for Semantic Visual Navigation

Organizers: Alexander Toshev
 Anelia Angelova
 Niko Sünderhauf
 Ronald Clark
 Andrew Davison

Location: 103B
Time: 0830-1645 (Full Day)

Summary: Visual navigation, the ability of an autonomous agent to find its way in a large, visually complex environment based on visual information, is a fundamental problem in computer vision and robotics.



This workshop provides a forum to promising ideas proposing to advance visual navigation by combining recent developments in deep and reinforcement learning. A special focus lies on approaches that incorporate more semantic information into navigation, and combine visual input with other modalities such as language.

Vision With Biased or Scarce Data

Organizers: Jan Ernst
 Ziyan Wu
 Kuan-Chuan Peng
 Srikrishna Karanam

Location: Hyatt Seaview B

Time: 0830-1300 (Half Day - Morning)



Summary: With the ever increasing appetite for data in machine learning, we need to face the reality that for many applications, sufficient data may not be available. Even if raw data is plenty, quality labeled data may be scarce, and if it is not, then relevant labeled data for a particular objective function may not be sufficient. The latter is often the case in tail end of the distribution problems, such as recognizing in autonomous driving that a baby stroller is rolling on the street. The event is rare in training and testing data, but certainly highly critical for the objective function of personal and property damage. Even the performance evaluation of such a situation is challenging. One may stage experiments geared towards particular situations, but this is not a guarantee that the staging conforms to the natural distribution of events, and even if, then there are many tail ends in high dimensional distributions, that are by their nature hard to enumerate manually. Recently the issue has been recognized more widely: DARPA for instance announced the program of Learning with Less Labels, that aims to reduce the number of labels required by a million-fold across a wide set of problems, vision included. In addition, there is mounting evidence of societal effects of data-driven bias in artificial intelligence such as in hiring and policy making with implications for non-government organizations as well as corporate social responsibility and governance. In this second workshop we would like to achieve two goals: (1) Raise awareness by having experts from academia, government and industry share about their perspectives, including on the impact of discriminatory biases of AI, and (2) Share the latest and greatest about biased and scarce data problems and solutions by distinguished speakers from academia and industry.

Landmark Recognition

Organizers: Bohyung Han Ondrej Chum
 Andre Araujo Torsten Sattler
 Bingyi Cao Giorgos Tolias
 Shih-Fu Chang Tobias Weyand
 Jack Sim Xu Zhang

Location: 104A

Time: 0845-1235 (Half Day - Morning)



Summary: This workshop fosters research on image retrieval and landmark recognition by introducing a novel large-scale dataset, together with evaluation protocols. In support of this goal, this year we are releasing Google-Landmarks-v2, a completely new, even larger landmark recognition dataset that includes over 5 million images (2x that of the first release) of more than 200 thousand different landmarks (an increase of 7x). Due to the difference in scale, this dataset is much more diverse and creates even greater challenges for state-of-the-art instance recognition approaches. Based on this new dataset, we are also announcing two new Kaggle challenges—Landmark Recognition 2019 and Landmark Retrieval 2019—and releasing the source code and model for Detect-to-Retrieve, a novel image representation suitable for retrieval of specific object instances.

Image Matching: Local Features and Beyond

Organizers: Vassileios Balntas
 Vincent Lepetit
 Johannes Schönberger
 Eduard Trulls
 Kwang Moo Yi

Location: Hyatt Regency A

Time: TBA (Half Day - Morning)



Summary: Traditional, keypoint-based formulations for image matching are still very competitive on tasks such as Structure from Motion (SfM), despite recent efforts on tackling pose estimation with dense networks. In practice, many state-of-the-art pipelines still rely on methods that stood the test of time, such as SIFT or RANSAC.

In this workshop, we aim to encourage novel strategies for image matching that deviate from and advance traditional formulations, with a focus on large-scale, wide-baseline matching for 3D reconstruction or pose estimation. This can be achieved by applying new technologies to sparse feature matching, or doing away with keypoints and descriptors entirely.

Detecting Objects in Aerial Images

Organizers: Gui-Song Xia Mihai Dactu
 Xiang Bai Marcello Pelillo
 Serge Belongie Liangpei Zhang
 Jiebo Luo

Location: Hyatt Regency F

Time: TBA (Half Day - Morning)



Summary: Object detection in Earth Vision, also known as Earth Observation and Remote Sensing, refers to the problem of localizing objects of interest (e.g., vehicles, airplanes and buildings) on the earth’s surface and predicting their corresponding categories. Observing plenty of instances from the overhead view provide a new way to understand the world. This is a relatively new field, with many new applications waiting to be developed. For movable categories, such as vehicles, ships, and planes, the orientation estimation is important for tracking. The majority of computer vision research focuses mostly on images from everyday life. However, the aerial imagery is a rich and structured source of information, yet, it is less investigated than it should be deserved. The task of object detection in aerial images is distinguished from the conventional object detection task in the following respects:

- The scale variations of object instances in aerial images are considerably huge.
- Many small object instances are densely distributed in aerial images, for example, the ships in a harbor and the vehicles in a parking lot.
- Objects in aerial images often appear in arbitrary orientations.

This workshop aims to draw attention from a wide range of communities and calls for more future research and efforts on the problems of object detection in aerial images. The workshop also contains a challenge on object detection in aerial images that features a new large-scale annotated image database of objects in aerial images, updated from DOTA-v1.0.

Understanding Subjective Attributes of Data: Focus on Fashion and Subjective Search

Organizers: Xavier Alameda-Pineda
 Miriam Redi
 Diane Larlus
 Kristen Grauman
 Nicu Sebe
 Shih-Fu Chang



Location: Hyatt Seaview C
Time: 0830-1230 (Half Day - Morning)

Summary: This workshop has a specific Focus on Fashion and Subjective Search. Indeed, fashion is influenced by subjective perceptions as well as societal trends, thus encompassing many subjective attributes (both individual and collective). Fashion is therefore a very relevant application for research on subjective understanding of data, and at the same time has great economic and societal impact. Moreover one of the hardest associated tasks is how to perform retrieval (and thus search) of visual content based on subjective attributes of data.

The automatic analysis and understanding of fashion in computer vision has growing interest, with direct applications on marketing, advertisement, but also as a social phenomena and in relation with social media and trending. Exemplar tasks are, for instance, the creation of capsules wardrobes. More fundamental studies address the design of unsupervised techniques to learn a visual embedding that is guided by the fashion style. The task of fashion artifact/landmark localization has also been addressed, jointly with the creation of a large-scale dataset. Another research line consists on learning visual representations for visual fashion search. The effect of social media tags on the training of deep architecture for image search and retrieval has also been investigated.

Visual Odometry and Computer Vision Applications Based on Location Clues

Organizers: Guoyu Lu
 Friedrich Fraundorfer
 Yan Yan
 Nicu Sebe
 Chandra Kambhamettu



Location: Hyatt Shoreline B
Time: 0830-1230 (Half Day - Morning)

Summary: Visual odometry has attracted substantial interest in computer vision, robotics and mechanical engineering communities, to name a few. Visual odometry estimates the location and orientation of a vehicle, robot and human. With the advent of autonomous driving and augmented reality, the applications of visual odometry are significantly growing. The development of smart-phones and cameras is also making the visual odometry more accessible to common users in daily life. With the increasing efforts devoted to accurately computing the position information, emerging applications based on location context, such as scene understanding, city navigation and tourist recommendation, have gained significant growth. The location information can bring a rich context to facilitate a large number of challenging problems, such as landmark and traffic sign recognition under various weather and light conditions, and computer vision applications on entertainment based on location information, such as Pokemon. The workshop publishes scalable algorithms and systems for addressing the ever increasing demands of accurate and real-time visual odometry, as well as the methods and applications based on the location clues.

Multimodal Learning and Applications

Organizers: Pietro Morerio
 Paolo Rota
 Michael Ying Yang
 Bodo Rosenhahn
 Vittorio Murino



Location: Seaside 7
Time: 0820-1300 (Half Day - Morning)

Summary: The exploitation of the power of big data in the last few years led to a big step forward in many applications of Computer Vision. However, most of the tasks tackled so far are involving mainly visual modality due to the unbalanced number of labelled samples available among modalities (e.g., there are many huge labelled datasets for images while not as many for audio or IMU based classification), resulting in a huge gap in performance when algorithms are trained separately.

This workshop aims to bring together communities of machine learning and multimodal data fusion. We expect contributions involving video, audio, depth, IR, IMU, laser, text, drawings, synthetic, etc. Position papers with feasibility studies and cross-modality issues with highly applicative flair are also encouraged therefore we expect a positive response from academic and industrial communities.

Explainable AI

Organizers: Quanshi Zhang
 Lixin Fan
 Bolei Zhou
 Sinisa Todorovic
 Tianfu Wu
 Ying Nian Wu



Location: Hyatt Beacon A
Time: 0800-1230 (Half Day - Morning)

Summary: Deep neural networks (DNNs) have no doubt brought great successes to a wide range of applications in computer vision, computational linguistics and AI. However, foundational principles underlying the DNNs' success and their resilience to adversarial attacks are still largely missing. Interpreting and theorizing the internal mechanisms of DNNs becomes a compelling yet controversial topic. The statistical methods and rule-based methods for network interpretation have much to offer in semantically disentangling inference patterns inside DNNs and quantitatively explaining the decisions made by DNNs. Rethinking DNNs explicitly toward building explainable systems from scratch is another interesting topic, including new neural architectures, new parameter estimation methods, new training protocols, and new interpretability-sensitive loss functions.

This workshop aims to bring together researchers, engineers as well as industrial practitioners, who concern about interpretability, safety, and reliability of artificial intelligence. Joint force efforts along this direction are expected to open the black box of DNNs and, ultimately, to bridge the gap between connectionism and symbolism of AI research. The main theme of the workshop is therefore to build up consensus on a variety of topics including motivations, typical methodologies, and prospective innovations of transparent and trustworthy AI. Research outcomes are also expected to have profound influences on critical industrial applications such as medical diagnosis, finance, and autonomous driving.

Towards Causal, Explainable and Universal Medical Visual Diagnosis

Organizers: Xiaodan Liang Eric Xing
 Christy Yuan Li Lawrence Carin
 Hao Wang Ricardo Henao



Location: Hyatt Beacon B
Time: 0830-1230 (Half Day - Morning)

Summary: Medical visual diagnosis has been gaining increased interest, and widely recognized in both academia and industry as an area of high impact and potential. In addition to classical problems such as medical image segmentation, abnormality detection and personalized diagnosis that benefits from the combination of deep learning approaches and big data, a more challenging goal towards causal, explainable and universal medical visual diagnosis has been urged by the recent availability of large-scale medical data and realistic industrial need. Specifically, medical decisions are usually made by collective analysis of multiple sources such as images, clinical notes, and lab tests, as well as combined human intelligence empowered by medical literature and domain knowledge. Having a single data-driven decision is insufficient for interactive assistance in clinical setting; a wider explanation on how and why the decision is made, and a deeper rationality on whether it can be justified by medical domain knowledge and personalized patient disease evolution is desired and necessary. Furthermore, this human-like intelligence can be more thoroughly explored in the recent surge of multi-modal tasks such as single-image and time-series medical image report generation, medical relational and casualty learning, and reinforcement learning for robust and unified diagnostic systems. The goal of this workshop is to allow researchers from machine learning, medical healthcare, and other disciplines to exchange ideas, advance an integrative reconciliation between theoretical analysis and industrial landing, and potentially shape the future of this area.

Energy Efficient Machine Learning and Cognitive Computing for Embedded Applications

Organizers: Raj Parihar Tao (Terry) Sheng
 Michael Goldfarb Krishna Nagar
 Satyam Srivastava Debu Pal
 Mahdi N. Bojnordi



Location: Hyatt Shoreline A
Time: 0800-1230 (Half Day - Morning)

Summary: As artificial intelligence and other forms of cognitive computing continue to proliferate into new domains, many new forums for dialogue and knowledge sharing have emerged. In the proposed workshop, the primary focus is on the discussion and dialogues on energy efficient techniques for cognitive computing and machine learning particularly for embedded applications and systems. For such resource constrained environments, performance alone is never sufficient, requiring system designers to balance performance with power, energy, and area (overall PPA metric).

The goal of this workshop is to provide a forum for researchers who are exploring novel ideas in the field of energy efficient machine learning and artificial intelligence for embedded applications. We also hope to provide a solid platform for forging relationships and exchange of ideas between the industry and the academic world through discussions and active collaborations.

Women in Computer Vision

Organizers: Irene Amerini
 Elena Balashova
 Sayna Ebrahimi
 Kathryn Leonard
 Arsha Nagrani
 Amaia Salvador



Location: Hyatt Regency A
Time: 1330-1800 (Half Day - Afternoon)

Summary: Computer vision has become one of the largest computer science research communities. We have made tremendous progress in recent years over a wide range of areas, including object recognition, image understanding, video analysis, 3D reconstruction, etc.

However, despite the expansion of our field, the percentage of female faculty members and researchers both in academia and in industry is still relatively low. As a result, many female researchers working in computer vision may feel isolated and do not have a lot of opportunities to meet with other women.

The goals of this workshop are to:

- Raise visibility of female computer vision researchers by presenting invited research talks by women who are role models in this field.
- Give opportunities to junior female students or researchers to present their work via a poster session and travel awards.
- Share experience and career advice for female students and professionals.

The half-day workshop on Women in Computer Vision is a gathering for both women and men working in computer vision. Researchers at all levels who are interested in computer vision are welcome and encouraged to attend the workshop. Travel grants will be offered to selected female presenters of oral and poster sessions.

ScanNet Indoor Scene Understanding

Organizers: Angela Dai
 Angel X. Chang
 Manolis Savva
 Matthias Niessner



Location: Hyatt Seaview B
Time: 1345-1730 (Half Day - Afternoon)

Summary: 3D scene understanding for indoor environments is becoming an increasingly important area. Application domains such as augmented and virtual reality, computational photography, interior design, and autonomous mobile robots all require a deep understanding of 3D interior spaces, the semantics of objects that are present, and their relative configurations in 3D space.

We present the first comprehensive challenge for 3D scene understanding of entire rooms at the object instance-level with 5 tasks based on the ScanNet dataset. The ScanNet dataset is a large-scale semantically annotated dataset of 3D mesh reconstructions of interior spaces (approx. 1500 rooms and 2.5 million RGB-D frames). It is used by more than 480 research groups to develop and benchmark state-of-the-art approaches in semantic scene understanding. A key goal of this challenge is to compare state-of-the-art approaches operating on image data (including RGB-D) with approaches operating directly on 3D data (point cloud, or surface mesh representations). Additionally, we pose both object category label prediction (commonly referred to as semantic segmentation), and instance-level object recognition (object instance prediction and category label prediction).

Computer Vision for UAVs

Organizers: Kristof Van Beeck
Toon Goedemé
Tinne Tuytelaars
Davide Scaramuzza
Marian Verhelst



Location: Hyatt Seaview A

Time: 1330-1800 (Half Day - Afternoon)

Summary: The UAVision2019 workshop focuses on state-of-the-art real-time image processing on-board of Unmanned Aerial Vehicles. Cameras are ideal sensors for drones as they are lightweight, power-efficient and an enormously rich source of information about the environment in numerous applications. Although lots of information can be derived from camera images using the newest computer vision algorithms, the use of them on-board of UAVs poses unique challenges. Remote processing is possible, although this requires a wireless link with high bandwidth, minimal latency and an ultra-reliable connection. Truly autonomous drones should not have to rely on a wireless data link, thus on-board real-time processing is a necessity. Because of the limitations of UAVs (lightweight processing devices, limited on-board computational power, limited electrical power), extreme algorithmic optimization and deployment on state-of-the-art embedded hardware (e.g. embedded GPUs) is the only solution. In this workshop we focus on enabling embedded processing on drones, making efficient use of specific embedded hardware and highly optimizing computer vision algorithms towards real-time applications. Apart from the regular paper submissions track we also organize an ERTI (Embedded Real-Time Inference) challenge. In this challenge competitors need to develop a pedestrian detection framework which runs on an NVIDIA Jetson TX2 with a minimum processing speed of 5 FPS.

Compact and Efficient Feature Representation and Learning in Computer Vision

Organizers: Li Liu
Yulan Guo
Wanli Ouyang
Jiwen Lu
Matti Pietikäinen
Luc Van Gool



Location: Hyatt Shoreline B

Time: 1350-1800 (Half Day - Afternoon)

Summary: Classification networks have been dominant in visual recognition, from image-level classification to region-level classification (object detection) and pixel-level classification (semantic segmentation, human pose estimation, and facial landmark detection). We argue that the classification network, formed by connecting high-to-low convolutions in series, is not a good choice for region-level and pixel-level classification because it only leads to rich low-resolution representations or poor high-resolution representations obtained with upsampling processes.

We propose a high-resolution network (HRNet). The HRNet maintains high-resolution representations by connecting high-to-low resolution convolutions in parallel and strengthens high-resolution representations by repeatedly performing multi-scale fusions across parallel convolutions. We demonstrate the effectiveness on pixel-level classification, region-level classification, and image-level classification. The HRNet turns out to be a strong replacement of classification networks (e.g., ResNets, VGGNets) for visual recognition.

Computer Vision After 5 Years

Organizers: Deepak Pathak
Shubham Tulsiani
Saurabh Gupta
Abhinav Gupta



Location: 104A

Time: TBA (Half Day - Afternoon)

Summary: There has been rapid progress in the field of Computer Vision in recent times, both in terms of solutions to established problems and emergence of new areas. At such a pace, it is increasingly difficult, but all the more important, for individual researchers to take stock of the changing landscape and be thoughtful about research directions being pursued. The goal of this workshop is to bring together pioneers in the field and ask them to predict the direction of our field in the next five years -- what in their view would be areas where progress would be made, or important problems that would remain open. We hope that this should not only be an exciting topic of discussion, but also provide seedling ideas to young graduate students who are the driving force of research in our community.

Benchmarking Multi-Target Tracking: How Crowded Can It Get?

Organizers: Laura Leal-Taixe
Hamid Rezaatofighi
Anton Milan
Javen Shi
Konrad Schindler

Patrick Dendorfer
Daniel Cremers
Stefan Roth
Ian Reid



Location: Hyatt Beacon B

Time: 1330-1800 (Half Day - Afternoon)

Summary: For this 4th edition of our Benchmarking Multi-Target Tracking (MOTChallenge) Workshop, we want to push the limits of both detectors and trackers with the introduction of our CVPR19 challenge. This dataset will be the foundation of the new MOT19 challenge that will be realized later this year. Following the strict annotation protocols of MOT16 and MOT17, we introduce 8 challenging and extremely crowded scenes, with over 160 pedestrians per frame on average. Despite the large number of objects in the videos, the sequences are longer and reach up to 3300 frames. We created two special CVPR challenges for tracking and detection, and accepted new submissions ahead of the workshop. This workshop will analyze the performance of state-of-the-art detectors and trackers on these very crowded scenes and discuss limitations of current methods with increasing pedestrian density. The best challenge submissions will have the opportunity to present their model during a poster session. In addition, we have invited six experts (Georgia Gkioxari, Paul Voigtlaender, Silvio Savarese, Rita Cucchiara, Jim Rehg, Ming-Hsuan Yang) to speak about different methods, challenges, and novelties for multi-object tracking. Furthermore, the workshop should encourage a dedicated discussion among participants on how to improve multi-target tracking evaluation and ideas on how to expand the current benchmark. These discussions in previous editions of the workshop have helped us tremendously in shaping MOTChallenge and significantly contributed to creating a widely used and perhaps the most popular multi-object tracking benchmark in the community.

3D Computer Vision in Medical Environment

Organizers: Vivek Singh
Yao-jen Chang
Ankur Kapoor



Location: Hyatt Beacon A

Time: 1330-1800 (Half Day - Afternoon)

Summary: Over the years, progress on computer vision research has effectively benefitted medical domain, leading to development of several high impact image-guided interventions and therapies. While the past couple of years have seen tremendous progress on 3D computer vision, especially in ADAS or driver-less navigation domains, the impact on medical domain has been limited. This workshop is to bring together the practitioners of 3D computer vision and medical domain and engage in a dialogue emphasizing the key recent advances and bridging the gap between 3d computer vision research and potential medical applications.

The primary topics will include, but not limited to the following,

- 3D human body modeling and estimation Human body modeling (or patient modeling) is critical to several applications such as patient positioning for medical scanning, support for prosthetic design, computed assisted rehabilitation systems.
- Non-rigid shape representation For human body, organs, vessels etc.; topics that emphasize the trade offs involved in volumetric or point-based representations
- Endoscopic imaging and analysis for surgical guidance 3D reconstruction using endoscopic imaging to provide guidance to surgical procedures
- Scene representation and modeling for surgical and scanning workflow analysis To localize the physicians and medical devices for workflow analysis during medical scanning as well as surgery and/or enable augmented reality applications
- Scene reconstruction for navigation For navigation and path planning for devices

Conceptual Captions Challenge

Organizers: Radu Soricut
Bohyung Han
Mohit Bansal
Yoav Artzi
Leonid Sigal
Ting Yao



Location: Seaside 7

Time: 1330-1720 (Half Day - Afternoon)

Summary: Automatic caption generation is the task of producing a natural-language utterance (usually a sentence) that describes the visual content of an image. One of the most critical limitations is limited understanding of the complex evaluation problem. The goal of this workshop is two-fold: (a) coalescing community effort around a new challenging web-scale image-captioning dataset, and (b) formalizing a human evaluation protocol, which is expected to boost both evaluation reliability and efforts on automatic quality estimation of caption generation. The dataset consisting of ~3.3 million image/caption pairs is publicly available (Conceptual Captions). We will employ a protocol to accurately estimate the quality of the image captions generated by the challenge participants, using both automatic metrics and human evaluators.

Beyond better understanding of the current state-of-the-art, our evaluation will allow us to observe correlation or discrepancy between automatic and human evaluation metrics. This will provide

additional support for the creation of new automatic evaluation metrics that better reflect human judgments. In addition, we plan to release the resulting human judgments on caption quality (for a subset of the test set containing images with appropriate licence rights), with the goal of providing additional data for improving algorithmic methods for caption quality assessment in the absence of groundtruth captions.

Mutual Benefits of Cognitive and Computer Vision

Organizers: Minh Hoai Nguyen
Krista A. Ehinger
Dimitris Samaras
Gregory Zelinsky



Location: Hyatt Shoreline A

Time: 1330-1800 (Half Day - Afternoon)

Summary: State-of-the-art computer vision systems have benefited greatly from our understanding of the human visual system, and research in human vision has benefited from image processing and modeling techniques from computer vision. Current advances in machine learning (e.g., deep learning) have produced computer vision systems which rival human performance in various narrowly-defined tasks such as object classification. However, the biological vision system remains the gold standard for efficient, flexible, and accurate performance across a wide range of complex real-world tasks. We believe that close collaboration between the fields of human and computer vision will lead to further breakthroughs in both fields. The goal of this workshop is to investigate the relationships between biological and computer vision, and how we can use insights from one to better understand the other.

Our workshop will broadly address relationships between biological vision and computer vision, but questions of particular interest include: 1) How does the concept of "attention" in computer vision relate to processing in the human visual attention system? 2) How do the features used by humans to represent objects and scenes compare to the features learned by artificial deep networks to perform large-scale image classification? And 3) Should computer vision models be designed after the primate visual system? This workshop aims to foster a level of understanding that is more than the sum of its questions, one where unifying principles emerge that shape new questions and new directions for seeking answers.

Perception Beyond the Visible Spectrum

Organizers: Riad Hammoud
Michael Teutsch
Angel D. Sappa
Yi Ding



Location: Hyatt Seaview C

Time: 1330-1800 (Half Day - Afternoon)

Summary: Since its inception in 2004, the Perception Beyond the Visible Spectrum workshop series (IEEE PBVS) has been one of the key events in the computer vision and pattern recognition (CVPR) community featuring imaging, sensing and exploitation algorithms in the non-visible spectrum (infrared, thermal, radar, SAR, millimeters wave, LiDAR, ...) for various applications including autonomous driving, aerial robotics, remote sensing, surveillance, and medical applications. This year PBVS hosts three keynote speakers from NASA, FLIR, and UCB/Stanford, and includes 23 papers in the program. The best paper award is sponsored by TuSimple.

Monday, June 17

NOTE: Use the QR code for each tutorial's website for more information on that tutorial. Here's the QR code to the CVPR Tutorials page.



0730-1730 Registration (Promenade Atrium & Plaza)

0730-0900 Breakfast (Pacific Ballroom)

1000-1045 Morning Break (Exhibit Hall)

1200-1330 Lunch (Pacific Ballroom)

1515-1600 Afternoon Break (Exhibit Hall)

Tutorial: Learning-Based Depth Estimation From Stereo and Monocular Images: Successes, Limitations and Future Challenges

Organizers: Matteo Poggi
Fabio Tosi
Konstantinos Batsos
Philippos Mordohai
Stefano Mattoccia

Location: 204

Time: Half Day - Morning (0900-1230)

Description: Obtaining dense and accurate depth measurement is of paramount importance for many 3D computer vision applications. Stereo matching has undergone a paradigm shift in the last few years due to the introduction of learning-based methods that replaced heuristics and hand-crafted rules. While in early 2012 the KITTI dataset highlighted how stereo matching was still an open problem, the recent success of Convolutional Neural Networks has led to tremendous progress and has established these methods as the undisputed state of the art. Similar observations can be made on all recent benchmarks, such as the KITTI 2012 and 2015, the Middlebury 2014 and the ETH3D benchmark, the leaderboards of which are dominated by learning-based methods.

The tutorial will cover conventional and deep learning methods that have replaced the components of the conventional stereo matching pipeline, end-to-end stereo systems and confidence estimation. The second part will focus on related problems, specifically single-view depth estimation, that have also benefited from the availability of ground truth datasets and learning algorithms. The tutorial will conclude with open problems including generalization as well as unsupervised and weakly supervised training.



Tutorial: Apollo: Open Autonomous Driving Platform

Organizers: Tae Eun Choe Jiangtao Hu
Liang Wang Ruigang Yang
Shiyu Song Jaewon Jung

Location: 203B

Time: Half Day - Morning (0900-1230)

Description: Apollo is the largest open autonomous driving platform with a full stack of H/W and S/W developed by the autonomous driving community. We will present the ongoing research in Apollo and discuss about future direction of autonomous driving. We will mainly discuss 5 topics: perception, simulation, sensor fusion, localization, and control: 1) Perception: we will review pros and cons of each sensor and discuss what functionality and level of autonomy can be achieved with such sensors. We will also discuss the main issues to reach L4 autonomy with cameras. 2) Simulation: we will demonstrate game-engine based simulation for training and evaluation of perception algorithms using camera and lidar sensors. 3) Sensor fusion: we will present how to learn a prior and belief function of each sensor and fuse all sensor output using Dempster-Shafer theory. 4) Localization: we will present automated HD map generation in a large scale and show a highly precise localization algorithm integrating GNSS, IMU, camera, and lidar. 5) Control: We will also explain how to perform multiple iteration optimization in planning and introduce a learning-based dynamic control modeling and its application in simulation.



Tutorial: Textures, Objects, Scenes: From Handcrafted Features to CNNs and Beyond

Organizers: Li Liu
Bolei Zhou
Liang Zheng
Wanli Ouyang

Location: 104C

Time: Half Day - Morning (0900-1230)

Description: This tutorial aims to review computer vision techniques before and after the deep learning era, in critical domains such as object detection, texture classification, scene understanding and instance retrieval. In the computer vision community, dramatic evolution is witnessed in the past 25 years, especially in visual recognition. The tremendous success cannot be made possible without the development of feature representation and learning approaches, which are at the core of many visual recognition problems such as texture recognition, image classification, object detection and recognition, scene classification and content based instance retrieval. In specific, we will focus on four closely related visual recognition problems at different levels: texture recognition, objects detection and recognition, scene understanding, and content based image retrieval. These problems have received significant attention from both academia and industry in the field of computer vision and pattern recognition. For each problem, this tutorial will firstly review the milestones in the two development stages, then present an overview of the current frontier and state of the art performance on leading benchmark datasets, and finally discuss the possible future research directions.



Tutorial: OpenCV 4.x and More New Tools for CV R&D

Organizers: Alexander Bovyryn
Vadim Pisarevsky
Nikita Manovich

Location: 202B

Time: Half Day - Afternoon (1330-1730)

Description: Tutorial on the new OpenCV 4.x features and exciting tools from "github.com/opencv" tool set. It covers OpenCV 4.0 features introduction, deep learning module usage with code samples in C++, Python, Java and JavaScript. There will also be a practical hands-on session where participants will play with the new functionality. In particular participants will know:

- How to run deep networks on Android device with OpenCV 4.0
- How to run deep networks in a browser with OpenCV 4.0
- Custom deep learning layers support in OpenCV 4.0

OpenCV now hosts Computer Vision Annotation Tool (CVAT) which is web-based, free, online, interactive video and image annotation tool for computer vision. There will be practical session on CVAT.

We will also provide update on Open Model Zoo and discuss CNN optimization tools.

**Tutorial: Representing Cause-and-Effect in a Tensor Framework**

Organizers: M. Alex O. Vasilescu
Jean Kossaifi
Lieven DeLathauwer

Location: 203C

Time: Half Day - Afternoon (1300-1715)

Description: Most observable data are multimodal and the result of several causal factors of data formation. Similarly, natural images are the compositional consequence of multiple causal factors related to scene structure, illumination, and imaging. Tensor algebra, the algebra of higher-order tensors offers a potent mathematical framework for explicitly representing and disentangling the causal factors of data formation. Theoretical evidence has shown that deep learning is a neural network equivalent to multilinear tensor decomposition, while a shallow network corresponds to linear tensor factorization (aka. CANDECOMP/Parafac tensor factorization).

There are two main classes of tensor decompositions which generalize different concepts of the matrix SVD, linear decomposition (rank-K decomposition) and multilinear decomposition (orthonormal matrices), which we will address, in addition to various tensor factorizations under different constraints.

We will also discuss several multilinear representations, Multilinear PCA, Multilinear ICA (which should not be confused to the computation of the linear ICA basis vectors by employing the CP tensor decomposition on a tensor that contains the higher order statistics of a data matrix), Compositional Hierarchical Tensor Factorization, Block Tensor Decomposition, etc. and introduce the multilinear projection operator, tensor pseudo-inverse and the identity tensor which are important in performing recognition in a tensor framework.

Tensor factorizations can also be efficiently combined with deep learning using TensorLy, a high level API for tensor algebra decomposition and regression. Deeply tensorized architecture results in state-of-the-art performance, large parameter savings and computational speed-ups on a wide range of applications.

**Tutorial: Deep Reinforcement Learning for Computer Vision**

Organizers: Jiwen Lu
Liangliang Ren
Yongming Rao

Location: 203A

Time: Half Day - Afternoon (1300-1700)

Description: In recent years, deep reinforcement learning has been developed as a basic technique in machine learning and successfully applied to a wide range of computer vision tasks. This tutorial will overview the trend of deep reinforcement learning techniques and discuss how they are employed to boost the performance of various computer vision tasks. First, we briefly introduce the basic concept of deep reinforcement learning, and show the key challenges in different computer vision tasks. Second, we introduce some deep reinforcement learning techniques and their varieties for computer vision tasks: policy learning, attention-aware learning, non-differentiable optimization and multi-agent learning. Third, we present several applications of deep reinforcement learning in different fields of computer vision. Lastly, we will discuss some open problems in deep reinforcement learning to show how to further develop more advanced algorithms for computer vision in the future.

**Tutorial: Bringing Robots to the Computer Vision Community**

Organizers: Adithyavairavan Murali
Dhiraj Gandhi
Lerrel Pinto
Deepak Pathak
Saurabh Gupta

Location: 204

Time: Half Day - Afternoon (1330-1730)

Description: There's been a surge of interest in robotics in the vision community. Several recent papers tackle robotics problems: mobile navigation, visual servoing for manipulation and navigation, visual grasping/pushing, localization, embodied visual-question answering, vision and language navigation, mobility simulators. Several papers investigate how to effectively study data collected from robots for visual learning. Research also focuses on video analysis for learning affordances. Active vision is also becoming increasingly popular.

While these works demonstrate impressive results, most of them shy away from showing results on real robots, likely because of the lack of expertise in the community, inaccessibility to robotic platforms, and even just the fear of dealing with robotic hardware. This limits the impact of these works, as the robotics community typically believes in results only when they see successful deployments on physical platforms. Furthermore, abstracting out the physical system, also removes important and interesting research challenges. Thus, the exposure to physical robots can guide practitioners toward more fruitful research directions and lead to more impactful work.

This tutorial's goal is to fill this gap in expertise and equip interested participants with basic tools that are useful for building, programming and operating robots. We'll focus on popular use cases in the community, and use a running example on an open-source low-cost manipulator with a mobile base. We believe this expertise will enable computer vision researchers to better understand perception issues with real robots, demonstrate their algorithms on real systems in the real world, and make it easier for research to transfer between vision and robotics communities.



Monday, June 17

NOTE: Use the QR code for each workshop's website to find the workshop's schedule. Here's the QR code to the CVPR Workshops page.



0730-1730 Registration (Promenade Atrium & Plaza)

0730-0900 Breakfast (Pacific Ballroom)

1000-1045 Morning Break (Exhibit Hall; Hyatt Foyer)

1200-1330 Lunch (Pacific Ballroom)

1515-1600 Afternoon Break (Exhibit Hall; Hyatt Foyer)

Egocentric Perception, Interaction and Computing

Organizers: Dima Damen
Antonino Furnari
Walterio Mayol-Cuevas
David Crandall
Giovani Maria Farinella
Kristen Grauman

Location: Seaside 7

Time: 0900-1800 (Full Day)

Summary: Egocentric perception introduces challenging and fundamental questions for computer vision as motion, real-time responsiveness, and generally uncontrolled interactions in the wild, are more frequently encountered. Questions such as what to interpret as well as what to ignore, how to efficiently represent actions, and how captured information can be turned into useful data for guidance or log summaries become central. Importantly, eyewear devices are becoming increasingly popular, both as research prototypes and off-the-shelf products. They capture interactions with the world, and enable a rich set of additional sources of information beyond images and videos. These include gaze information, audio, geolocation and IMU data.

The EPIC@X workshop series is dedicated to pushing the state of the art in research and methodologies in emerging research in egocentric perception, video computing with eyewear systems, multi-sensor responses, and egocentric interaction. An active worldwide mailing list for exchanging code, news, jobs and datasets brings together the community (epic-community@bristol.ac.uk).

In addition to demos and posters of ongoing and recently published works, EPIC@CVPR2019 has three-track challenges, evaluated on the largest egocentric dataset to date EPIC-KITCHENS - 11.5M images captured in people's native home environments. These are:

- Action Recognition
- Action Anticipation
- Active Object Recognition

Challenge winners and works that use it will be invited to present results as well as novel ideas to this edition of the workshop series.



Applications of Computer Vision and Pattern Recognition to Media Forensics

Organizers: Cristian Canton Scott McCloskey
Larry S. Davis Laura Leal-Taixé
Edward Delp Paul Natsev
Patrick Flynn Christoph Bregler

Location: 102A

Time: 0900-1800 (Full Day)

Summary: The recent advent of techniques that generate photo-realistic fully synthetic images and videos, and the increasing prevalence of misinformation associated with such fabricated media have raised the level of interest in the computer vision community. Both academia and industry have addressed this topic in the past, but only recently, with the emergence of more sophisticated ML and CV techniques, has multimedia forensics become a broad and prominent area of research.



Computer Vision for Microscopy Image Analysis

Organizers: Mei Chen
Dimitris Metaxas
Steve Finkbeiner
Daniel J. Hoepfner

Location: 102C

Time: 0830-1800 (Full Day)

Summary: High-throughput microscopy enables researchers to acquire thousands of images automatically over a matter of hours. This makes it possible to conduct large-scale, image-based experiments for biological discovery. The main challenge and bottleneck in such experiments is the conversion of "big visual data" into interpretable information and hence discoveries. Visual analysis of large-scale image data is a daunting task. Cells need to be located and their phenotype (e.g., shape) described. The behaviors of cell components, cells, or groups of cells need to be analyzed. The cell lineage needs to be traced. Not only do computers have more "stamina" than human annotators for such tasks, they also perform analysis that is more reproducible and less subjective. The post-acquisition component of high-throughput microscopy experiments calls for effective and efficient computer vision techniques.

This workshop intends to draw more visibility and interest to this challenging yet fruitful field and establish a platform to foster in-depth idea exchange and collaboration. We aim for broad scope, topics of interest include but are not limited to:

- Image acquisition
- Image calibration
- Background correction
- Object detection
- Segmentation
- Stitching and Registration
- Event detection
- Object tracking
- Shape analysis
- Texture analysis
- Classification
- Big image data to knowledge
- Image datasets and benchmarking



Visual Question Answering and Dialog

Organizers: Abhishek Das
Ayush Shrivastava
Karan Desai
Yash Goyal
Aishwarya Agrawal
Amanpreet Singh
Meet Shah
Drew Hudson
Satwik Kottur

Rishabh Jain
Vivek Natarajan
Stefan Lee
Peter Anderson
Xinlei Chen
Marcus Rohrbach
Dhruv Batra
Devi Parikh

Location: Seaside Ballroom B
Time: 0900-1800 (Full Day)

Summary: To further progress towards the grand goal of building agents that can see and talk, we are organizing the Visual Question Answering and Dialog Workshop. Its primary purposes are two-fold. First is to benchmark progress in these areas by hosting the Visual Question Answering (VQA) and Visual Dialog challenges. The VQA Challenge will have three tracks this year:

- 1) VQA 2.0 (<https://visualqa.org/challenge>): This track is the 4th challenge on the VQA v2.0 dataset introduced in Goyal et al., CVPR 2017.
- 2) TextVQA (<https://textvqa.org/challenge>): This track is the 1st challenge on the TextVQA dataset. TextVQA requires algorithms to read and reason about text in the image to answer a given question.
- 3) GQA (<https://cs.stanford.edu/people/dorarad/gqa/challenge.html>): This track is the 1st challenge on the GQA dataset. GQA is a new dataset that focuses on real-world compositional reasoning.

In addition, we will also be organizing the 2nd Visual Dialog Challenge (<https://visualldialog.org/challenge>). Visual Dialog requires an AI agent to hold a meaningful dialog with humans in natural, conversational language about visual content.

The second goal of this workshop is to continue to bring together researchers interested in visually-grounded question answering, dialog systems, and language in general to share state-of-the-art approaches, best practices, and future directions in multi-modal AI. In addition to an exciting lineup of invited talks, we invite submissions of extended abstracts describing work in vision + language + action.

Multi-Modal Learning From Videos

Organizers: Chuang Gan
Boqing Gong
Xiaolong Wang
Limin Wang

Location: 201B
Time: 0850-1700 (Full Day)

Summary: Video data is explosively growing as a result of ubiquitous acquisition capabilities. The videos captured by smart mobile phones, from ground surveillance, and by body-worn cameras can easily reach the scale of gigabytes per day. While the “big video data” is a great source for information discovery and extraction, the computational challenges are unparalleled. Intelligent algorithms for automatic video understanding, summarization, retrieval, etc. have emerged as a pressing need in such context. Progress on this topic will enable autonomous systems for quick and decisive acts based on the information in the videos, which otherwise would not be possible.



Deep Learning for Geometric Shape Understanding

Organizers: Ilke Demir
Kathryn Leonard
Géraldine Morin
Camilla Hahn

Location: Seaside 4
Time: 0900-1730 (Full Day)

Summary: Computer vision approaches have made tremendous efforts toward understanding shape from various data formats, especially since entering the deep learning era. Although accurate results have been obtained in detection, recognition, and segmentation, there is less attention and research on extracting topological and geometric information from shapes. These geometric representations provide compact and intuitive abstractions for modeling, synthesis, compression, matching, and analysis. Extracting such representations is significantly different from segmentation and recognition tasks, as they contain both local and global information about the shape.

Deep Learning for Geometric Shape Understanding Workshop aims to bring together researchers from computer vision, computer graphics, and mathematics to advance the state of the art in topological and geometric shape analysis using deep learning. The workshop also hosts The SkelNetOn Challenge, which is structured around shape understanding in three pixel, point, and parametric domains. We provide shape datasets, some complementary resources, and the evaluation platform for novel and existing approaches to compete. The winner of each track will receive a Titan RTX GPU.

Fine-Grained Visual Categorization

Organizers: Ryan Farrell
Oisin Mac Aodha
Subhransu Maji
Serge Belongie

Location: Seaside Ballroom A
Time: 0900-1730 (Full Day)

Summary: Fine categorization, i.e., the fine distinction into species of animals and plants, of car and motorcycle models, of architectural styles, etc., is one of the most interesting and useful open problems that the machine vision community is just beginning to address. Aspects of fine categorization (called “subordinate categorization” in the psychology literature) are discrimination of related categories, taxonomization, and discriminative vs. generative learning.

Fine categorization lies in the continuum between basic level categorization (object recognition) and identification of individuals (face recognition, biometrics). The visual distinctions between similar categories are often quite subtle and therefore difficult to address with today’s general-purpose object recognition machinery. It is likely that radical re-thinking of some of the matching and learning algorithms and models that are currently used for visual recognition will be needed to approach fine categorization.

This workshop will explore computational questions of modeling, learning, detection and localization. It is our hope that the invited talks, including researchers from psychology and psychophysics, will shed light on human expertise and human performance in subordinate categorization and taxonomization.



Computer Vision for AR/VR

Organizers: Sofien Bouaziz
Matt Uyttendaele
Andrew Rabinovich
Fernando De la Torre
Serge Belongie



Location: 202C
Time: TBA (Full Day)

Summary: Augmented and Virtual Reality (AR/VR) have been around for more than 30 years. These technologies found their initial applications in the military aircraft arena, with Pilot Head-Mounted Displays, flight simulations and later in entertainment and industry. However, it was only recently that these technologies have found their way into the commercial world mostly due to advances in mobile processing, new specialized AR-chipsets, reduced development cost, the ubiquity of wireless broadband connections, and new algorithmic developments in computer vision.

VR is more mature than existing AR technology and it has already showed some effective industry use-cases as well, from training (e.g. astronaut training, flight simulators), gaming and real estate applications to tourism, while mobile-driven AR effects such as virtual try-on of makeup, face masks or games such as Pokemon Go are the first validation of consumer AR use. Although AR is less mature than VR due to technology limitations, lack of standardization and a higher price tag, it is expected to be game-changer technology in areas like manufacturing, healthcare and logistics in the next few years once the technological limitations of the traditional headsets are resolved.

Both AR and VR technologies are expected to reach a mainstream adoption, comparable to the adoption of smartphones (one third of American households have three or more smartphones). In order to make this possible, Computer Vision plays a fundamental role in AR/VR systems to see, analyze and understand the world.

Autonomous Driving

Organizers: Fisher Yu
Jose M. Alvarez
Oscar Beijbom
John Leonard
Markus Enzweiler
Antonio M. Lopez
Tomas Pajdla
David Vazquez

Holger Caesar
Zhengping Che
Haofeng Chen
Haoping Bai
Guangyu Li
Tracy Li
Adrien Gaidon



Location: Terrace Theater
Time: 0830-1800 (Full Day)

Summary: The CVPR 2019 Workshop on Autonomous Driving (WAD) aims to gather researchers and engineers from academia and industry to discuss the latest advances in perception for autonomous driving. In this one-day workshop, we will have regular paper presentations, invited speakers, and technical benchmark challenges to present the current state of the art, as well as the limitations and future directions for computer vision in autonomous driving, arguably the most promising application of computer vision and AI in general. The previous chapters of the workshop at CVPR attracted hundreds of researchers to attend. This year, multiple industry sponsors also join our organizing efforts to push its success to a new level.

3D Scene Understanding for Vision, Graphics, and Robotics

Organizers: Siyuan Huang
Chuhang Zou
Hao Su
Alexander Schwing
Shuran Song

Siyuan Qi
Yixin Zhu
David Forsyth
Leonidas Guibas
Song-Chun Zhu

Location: Seaside 1
Time: 0900-1730 (Full Day)

Summary: Tremendous efforts have been devoted to 3D scene understanding over the last decade. Due to their success, a broad range of critical applications like 3D navigation, home robotics, and virtual/augmented reality have been made possible already, or are within reach. These applications have drawn the attention and increased aspirations of researchers from the field of computer vision, computer graphics, and robotics.

However, significantly more efforts are required to enable complex tasks like autonomous driving or home assistant robotics, which demand a deeper understanding of the environment compared to what is possible today. Such a requirement is because these complex tasks call for an understanding of 3D scenes across multiple levels, relying on the ability to accurately parse, reconstruct and interact with the physical 3D scene, as well as the ability to jointly recognize, reason and anticipate activities of agents within the scene. Therefore, 3D scene understanding problems become a bridge that connects vision, graphics and robotics research.

The goal of this workshop is to foster interdisciplinary communication of researchers working on 3D scene understanding (computer vision, computer graphics, and robotics) so that more attention of the broader community can be drawn to this field. Through this workshop, current progress and future directions will be discussed, and new ideas and discoveries in related fields are expected to emerge.

Safe Artificial Intelligence for Automated Driving

Organizers: Timo Sämam
Stefan Milz
Fabian Hüger
Peter Schlicht
Joachim Sicking

Stefan Rüping
Oliver Grau
Oliver Wasenmüller
Markus Enzweiler
Loren Schwarz

Location: 103A
Time: 0900-1715 (Full Day)

Summary: A conventional analytical procedure for the realization of highly automated driving reaches its limits in complex traffic situations. The switch to artificial intelligence is the logical consequence. The rise of deep learning methods is seen as a breakthrough in the field of artificial intelligence. A disadvantage of these methods is their opaque functionality, so that they resemble a black box solution. This aspect is largely neglected in the current research work and a pure increase in performance is aimed. The use of black box solutions represents an enormous risk in safety-critical applications such as highly automated driving. The development and evaluation of mechanisms that guarantee a safe artificial intelligence is required. The aim of this workshop is to increase the awareness of the active research area for this topic. The focus is on mechanisms that influence the deep learning model for computer vision in the training, test and inference phase.



SUMO Workshop: 360° Indoor Scene Understanding and Modeling

Organizers: Daniel Huber
Lyne Tchapmi
Frank Dellaert
Ilke Demir
Shuran Song
Rachel Luo

Location: 101B

Time: 0900-1730 (Full Day)



Summary: In computer vision, scene understanding and modeling encapsulate a diverse set of research problems, ranging from low-level geometric modeling (e.g., SLAM algorithms) to 3D room layout estimation. These tasks are often addressed separately, yielding only a limited understanding and representation of the underlying scene. In parallel, the popularity of 360° cameras has encouraged the digitization of the real world into augmented and virtual realities, enabling new applications such as virtual social interactions and semantically leveraged augmented reality.

The primary goals of the SUMO Challenge Workshop are:

- (i) To promote the development of comprehensive 3D scene understanding and modeling algorithms that create integrated scene representations (with geometry, appearance, semantics, and perceptual qualities).
- (ii) To foster research on the unique challenges of generating comprehensive digital representations from 360° imagery.

The workshop program includes keynotes from distinguished researchers in the field, oral presentations by the winners of the challenges, and a poster session with novel approaches in 3D scene understanding.

The SUMO Challenge, in conjunction with the workshop, provides a dataset and an evaluation platform to assess and compare scene understanding approaches that generate complete 3D representations with textured 3D models, pose, and semantics. The algorithms and datasets created for this competition may serve as reference benchmarks for future research in 3D scene understanding.

International Challenge on Activity Recognition

Organizers: Bernard Ghanem
Juan Carlos Niebles
Cees Snoek

Location: Grand Ballroom A

Time: 0900-1800 (Full Day)



Summary: This challenge is the 4th annual installment of International Challenge on Activity Recognition, previously called the ActivityNet Large-Scale Activity Recognition Challenge which was first hosted during CVPR 2016. It focuses on the recognition of daily life, high-level, goal-oriented activities from user-generated videos as those found in internet video portals.

We are proud to announce that this year's challenge will host a more diverse set of tasks which aim to push the limits of semantic visual understanding of videos as well as bridging visual content with human captions. These tasks range from large-scale activity classification from untrimmed videos, temporal and spatio-temporal activity localization, activity-based video captioning, active speaker localization, and analysis of instructional and surveillance videos.

Sight and Sound

Organizers: Andrew Owens
Jiajun Wu
William Freeman
Andrew Zisserman
Jean-Charles Bazin
Zhengyou Zhang
Antonio Torralba
Kristen Grauman

Location: Grand Ballroom B

Time: 0900-1700 (Full Day)

Summary: In recent years, there have been many advances in learning from visual and auditory data. While traditionally these modalities have been studied in isolation, researchers have increasingly been creating algorithms that learn from both modalities. This has produced many exciting developments in automatic lip-reading, multi-modal representation learning, and audio-visual action recognition.

Since pretty much every internet video has an audio track, the prospect of learning from paired audio-visual data — either with new forms of unsupervised learning, or by simply incorporating sound data into existing vision algorithms — is intuitively appealing, and this workshop will cover recent advances in this direction. But it will also touch on higher-level questions, such as what information sound conveys that vision doesn't, the merits of sound versus other “supplemental” modalities such as text and depth, and the relationship between visual motion and sound. We'll also discuss how these techniques are being used to create new audio-visual applications, such as in the fields of speech processing and video editing.



Target Re-Identification and Multi-Target Multi-Camera Tracking

Organizers: Ergys Ristani
Liang Zheng
Xiatian Zhu
Shiliang Zhang
Jingdong Wang
Shaogang Gong
Qi Tan
Carlo Tomasi
Richard Hartley

Location: 103C

Time: 0830-1500 (Full Day)

Summary: The 1st MTMCT and ReID workshop was successfully held at CVPR 2017. In the past two years, the MTMCT and REID community has been growing fast. As such, we are organizing this workshop for a second time, aiming to gather the state-of-the-art technologies and brainstorm future directions. We are especially welcoming ideas and contributions that embrace the relationship and future of MTMCT and ReID, two deeply connected domains. This workshop will encourage lively discussions on shaping future research directions for both academia and the industry. Examples of such directions are:

- How much do initial detections influence performance in MTMCT or ReID?
- How to improve the generalization ability of a MTMCT or ReID system?
- How and which ReID descriptors can be integrated into MTMCT systems?
- What can we learn by evaluating a MTMCT system in terms of ReID (and vice-versa)?
- How can ReID and MTMCT benefit each other?



EarthVision: Large Scale Computer Vision for Remote Sensing Imagery

Organizers: Devis Tuia
Jan Dirk Wegner
Ronny Hänsch
Bertrand Le Saux
Naoto Yokoya
Ilke Demir
Nathan Jacobs

HakJae Kim
Kris Koperski
Fabio Pacifici
Ramesh Raskar
Myron Brown
Mariko Burgin

Location: Hyatt Regency B
Time: 0830-1700 (Full Day)

Summary: Earth observation and remote sensing are ever-growing fields of investigation where computer vision, machine learning, and signal/image processing meet. The general objective is to provide large-scale, homogeneous information about processes occurring at the surface of the Earth exploiting data collected by airborne and spaceborne sensors. Earth Observation implies the need for multiple inference tasks, ranging from detection to registration, data mining, multisensor, multi-resolution, multi-temporal, and multi-modality fusion, to name just a few. It comprises ample applications like location-based services, online mapping services, large-scale surveillance, 3D urban modelling, navigation systems, natural hazard forecast and response, climate change monitoring, virtual habitat modelling, etc. The sheer amount of data needs highly automated workflows. This workshop aims at fostering collaboration between the computer vision and Earth Observation communities to boost automated interpretation of remotely sensed data and to raise awareness inside the vision community for this highly challenging and quickly evolving field of research with a big impact on human society, economy, industry, and the planet.



Dynamic Scene Reconstruction

Organizers: Armin Mustafa
Marco Volino
Michael Zollhoefer
Dan Casas
Adrian Hilton

Location: Hyatt Regency H
Time: 0900-1730 (Full Day)

Summary: Reconstruction of general dynamic scenes is motivated by potential applications in film and broadcast production together with the ultimate goal of automatic understanding of real-world scenes from distributed camera networks. With recent advances in hardware and the advent of virtual and augmented reality, dynamic scene reconstruction is being applied to more complex scenes with applications in Entertainment, Games, Film, Creative Industries and AR/VR/MR. This workshop aims to give an overview of the advances of computer vision algorithms in dynamic scene reconstruction to the target audience and will identify future challenges. The proposed workshop on Dynamic Scene Reconstruction is the first workshop to address the bottlenecks in this field by targeting the following objectives:

- Create a common portal for papers and datasets in dynamic scene reconstruction
- Discuss techniques to capture high-quality ground-truth for dynamic scenes
- Develop an evaluation framework for dynamic scene reconstruction



3D-WIDGET: Deep Generative Models for 3D Understanding

Organizers: David Vazquez
Sai Rajeswar
Pedro O. Pinheiro
Florian Golemo
Aaron Courville

Christopher Pal
Derek Nowrouzezahrai
Xavier Snelgrove
Fahim Mannan
Thomas Boquet

Location: Hyatt Regency A
Time: 0900-1720 (Full Day)

Summary: Understanding and modelling the 3D structure and properties of the real world is a fundamental problem in computer vision with broad applications in engineering, simulation, and sensing modalities in general. Humans have an innate ability to rapidly build models of the world that they are observing. With only few exposures to a new concept, humans relate and generalize this knowledge to other known concepts. And conversely, when visually taking in the environment, humans quickly identify concepts, geometry and the physical properties that are associated with them, even if parts of objects are obscured, for example objects being occluded in images.

This ability is crucial for building smarter AI systems. In these, the models have prior knowledge of objects that can occur in the world, learn to identify these objects and then make assumptions of how these objects would behave.

This workshop explores methods that learn 3D representations properly to be effectively used for learning in end-to-end differentiable settings such as using neural networks. This becomes especially important when we want to use 3D representations with other modalities like text, sound, images etc to capture certain important spatial properties that might be ignored otherwise. The goal of this workshop is to study how neural networks can learn to build such intricate 3D models of the world. Additionally, we are interested in research that explores and manipulates the 3D representation in the neural net.

ChaLearn Looking at People: Face Spoofing Attack

Organizers: Jun Wan
Sergio Escalera
Hugo Jair Escalante
Isabelle Guyon

Guodong Guo
Hailin Shi
Meysam Madadi
Shaopeng Tang

Location: 102B
Time: 0900-1710 (Full Day)

Summary: In the last ten years, face biometric research has been intensively studied by the computer vision community. Face recognition systems have been used in mobile, banking, and surveillance systems. For face recognition systems, face spoofing detection is a crucial stage which could affect security issues in government sectors. Although face anti-spoofing detection methods have been proposed so far, the problem is still unsolved due to the difficulty on the design of features and methods for spoof attacks. In the field of face spoofing detection, existing datasets are relatively small. This workshop introduces a new large-scale multimodal face spoofing attack dataset, the largest up-to-date in state-of-the-art for this purpose with an associated challenge.

For the proposed workshop we solicit submissions in all aspects of facial biometric systems and attacks. Most notably, the following are the main topics of interest:



- Novel methodologies on anti-spoofing detection in visual information systems
- Studies on novel attacks to biometric systems, and solutions
- Deep learning methods for biometric authentication systems using visual information
- Novel datasets and evaluation protocols on spoofing prevention on visual and multimodal biometric systems
- Methods for deception detection from visual and multimodal information
- Face antispoof attacks dataset (3D face Mask, multimodal)
- Deep analysis reviews on face antispoofing attacks
- Generative models (e.g. GAN) for spoofing attacks

Automated Analysis of Marine Video for Environmental Monitoring

Organizers: Derya Akkaynak
Anthony Hoogs
Suchendra Bahndarkar



Location: Seaside 3
Time: 0930-1700 (Full Day)

Summary: Manual annotation of imagery is the largest bottleneck to studying many important problems in the underwater domain. This workshop will convene experts and researchers from both marine science and computer vision communities to learn about the challenges, current work, and opportunities in marine imagery analysis.

Event-Based Vision and Smart Cameras

Organizers: Davide Scaramuzza
Guillermo Gallego
Kostas Daniilidis



Location: 101A
Time: 0800-1800 (Full Day)

Summary: This workshop is dedicated to event-based cameras, smart cameras (such as the SCAMP sensor) and algorithms. Event-based cameras are revolutionary vision sensors with three key advantages: a measurement rate that is almost 1 million times faster than standard cameras, a latency of microseconds, and a very high dynamic range. Because of these advantages, event-based cameras open frontiers that are unthinkable with standard cameras (which have been the main sensing technology for the past 60 years). These revolutionary sensors enable the design of a new class of algorithms to track a baseball in the moonlight, build a flying robot with the same agility of a fly, and perform structure from motion in challenging lighting and speed conditions. These sensors have covered the main news in recent years, with event-camera company Prophesee receiving \$40 million in investment from Intel and Bosch, and Samsung announcing mass production as well as its use in combination with IBM TrueNorth processor. Cellular processor arrays, such as the SCAMP sensor, are novel cameras in which each pixel has a programmable processor, thus they yield massively parallel processing near the image sensor. Because early vision computations are carried out entirely on-sensor, resulting vision systems have high speed and low-power consumption, enabling new embedded applications in areas such as robotics, AR/VR, automotive, surveillance, etc. This workshop will cover the sensing hardware, as well as the processing and learning methods needed to take advantage of the above-mentioned revolutionary cameras.

Robotic Vision Probabilistic Object Detection Challenge

Organizers: Niko Suenderhauf John Skinner
Feras Dayoub Haoyang Zhang
Anelia Angelova Gustavo Carneiro
David Hall

Location: 103B
Time: 0900-1700 (Full Day)

Summary: This workshop will bring together the participants of the first Robotic Vision Challenge, a new competition targeting both the computer vision and robotics communities.

The new challenge focuses on probabilistic object detection. The novelty is the probabilistic aspect for detection: A new metric evaluates both the spatial and semantic uncertainty of the object detector and segmentation system. Providing reliable uncertainty information is essential for robotics applications where actions triggered by erroneous but high-confidence perception can lead to catastrophic results.



Low-Power Image Recognition Challenge

Organizers: Yung-Hsiang Lu
Yiran Chen
Bo Chen
Alex Berg

Location: 201A
Time: 0930-1700 (Full Day)

Summary: This workshop will extend the successes of LPIRC in the past four years identifying the best vision solutions that can simultaneously achieve high accuracy in computer vision and energy efficiency. Since the first competition held in 2015, the winners' solutions have improved 24 times in the ratio of accuracy divided by energy.

LPIRC 2019 will continue the online + onsite options for participants. One online track with pre-selected hardware allows participants to submit their solutions multiple times without the need of traveling. One onsite track allows participants to bring their systems to CVPR.



New Trends in Image Restoration and Enhancement

Organizers: Radu Timofte Ming-Yu Liu
Shuhang Gu Zhiwu Huang
Lei Zhang Cosmin Ancuti
Ming-Hsuan Yang Codruta O. Ancuti
Luc Van Gool Jianrun Cai
Kyoung Mu Lee Seungjun Nah
Michael S. Brown Abdelrahman Abdelhamed
Eli Shechtman Richard Zhang

Location: 104A
Time: 0800-1800 (Full Day)

Summary: Image restoration and image enhancement are key computer vision tasks, aiming at the restoration of degraded image content, the filling in of missing information, or the needed transformation and/or manipulation to achieve a desired target (with respect to perceptual quality, contents, or performance of apps working on such images). Recent years have witnessed an increased interest from the vision and graphics communities in these fundamental topics of research. Not only has there been a constantly



growing flow of related papers, but also substantial progress has been achieved.

Each step forward eases the use of images by people or computers for the fulfillment of further tasks, as image restoration or enhancement serves as an important frontend. Not surprisingly then, there is an ever growing range of applications in fields such as surveillance, the automotive industry, electronics, remote sensing, or medical image analysis etc. The emergence and ubiquitous use of mobile and wearable devices offer another fertile ground for additional applications and faster methods.

This workshop aims to provide an overview of the new trends and advances in those areas. Moreover, it will offer an opportunity for academic and industrial attendees to interact and explore collaborations. Associated with the workshop are challenges on image dehazing, super-resolution, denoising, colorization, enhancement, and video deblurring and super-resolution.

Notes:

A large grid area for taking notes, consisting of approximately 20 columns and 30 rows of empty cells.

Dependable Deep Detectors: Verification and Certification of Neural Networks

Organizers: Shashank Pathak
Alessandro Abate
Armando Tacchella
Jelena Frtunikj

Location: Hyatt Seaview C

Time: 0830-1200 (Half Day - Morning)

Summary: Pulina et al. 2010 was one of the first works to look into verifying a multi-layer neural network. Though there have been several other methods proposed since then, adoption of these methods on state-of-the-art applications of deep learning has been missing. Much of this is also rooted in the fact that verifying even a piece-wise linear network is NP-hard and scalability is a huge challenge because most of the state-of-the-art neural networks are fairly deep. Note that this approach is slightly different from those finding effective strategies to counter the attacks such as recent papers from P. Liang; discussing such a connection is also a worthy pursuit. In contrast with major focuses of academic research in deep learning, provable safety is the single most important factor in the application of deep networks in any context that involves human and safety; a prominent example is in autonomous vehicles. Ability to certify such deep detectors are not only crucial to obtain a certified product that can be commercialized but can also save billions of dollars that will otherwise be lost in deploying a redundant, low-performance parallel system that is based on certifiable classical approaches. There are signs that industrial players of all sizes are deeply interested in exploring the avenues; wide spectrum partnerships across OEMs and tier-I are formed (e.g., industry-wide collaboration in Germany) and tier-II like Qualcomm, Nvidia and ARM have come up with hardware architecture ensuring redundancy-based safety.



Workshop & Challenge on Learned Image Compression

Organizers: George Toderici Johannes Ballé
Michele Covell Eirikur Agustsson
Wenzhe Shi Nick Johnston
Radu Timofte Fabian Mentzer
Lucas Theis

Location: 104B

Time: 0800-1230 (Half Day - Morning)

Summary: Our workshop aims to gather publications which will advance the field of image compression with and without neural networks.

Even with the long history of signal-processing oriented compression, taking new approaches to image processing have great potential, due to the proliferation of high-resolution cell-phone images and special hardware (e.g., GPUs). The potential in this area has already been demonstrated using recurrent neural networks, convolutional neural networks, and adversarial learning, many of these matching the best image-compression standards when measured on perceptual metrics. As such, we are interested in the various techniques associated with this class of methods. Broadly speaking, we would like to encourage the development of novel encoder/decoder architectures, novel ways to control information flow between the encoder and the decoder, and learn how to quantize (or learn to quantize) better.



Efficient Deep Learning for Computer Vision

Organizers: Peter Vajda
Pete Warden
Kurt Keutzer

Location: Hyatt Beacon A

Time: 0750-1230 (Half Day - Morning)



Summary: Computer Vision has a long history of academic research, and recent advances in deep learning have provided significant improvements in the ability to understand visual content. As a result of these research advances on problems such as object classification, object detection, and image segmentation, there has been a rapid increase in the adoption of Computer Vision in industry; however, mainstream Computer Vision research has given little consideration to speed or computation time, and even less to constraints such as power/energy, memory footprint and model size. Nevertheless, addressing all of these metrics is essential if advances in Computer Vision are going to be widely available on mobile devices. The morning session of the workshop goal is to create a venue for a consideration of this new generation of problems that arise as Computer Vision meets mobile system constraints. In the afternoon session, we will make sure we have a good balance between software, hardware and network optimizations with an emphasis on training of efficient neural networks with high performance computing architectures.

Bias Estimation in Face Analytics

Organizers: Rama Chellappa
Nalini Ratha
Rogerio Feris
Michele Merler
Vishal Patel

Location: Hyatt Seaview B

Time: 0840-1230 (Half Day - Morning)



Summary: Many publicly available face analytics datasets are responsible for great progress in face recognition. These datasets serve as sources of large amounts of training data as well as assessing performance of state-of-the-art competing algorithms. Performance saturation on such datasets has led the community to believe the face recognition and attribute estimation problems to be close to be solved, with various commercial offerings stemming from models trained on such data.

However, such datasets present significant biases in terms of both subjects and image quality, thus creating a significant gap between their distribution and the data coming from the real world. For example, many of the publicly available datasets underrepresent certain ethnic communities and over represent others. Most datasets are heavily skewed in age distribution. Many variations have been observed to impact face recognition including, pose, low-resolution, occlusion, age, expression, decorations and disguise. Systems based on a skewed training dataset are bound to produce skewed results. This mismatch has been evidenced in the significant drop in performance of state of the art models trained on those datasets when applied to images presenting lower resolution, poor illumination, or particular gender and/or ethnicity groups. It has been shown that such biases may have serious impacts on performance in challenging situations where the outcome is critical either for the subject or to a community. Often research evaluations are quite unaware of those issues, while focusing on saturating the performance on skewed datasets.

Fairness Accountability Transparency and Ethics in Computer Vision

Organizers: Timnit Gebru
Daniel Lim
Yabebal Fantaye
Margaret Mitchell
Anna Rohrbach

Location: Hyatt Seaview A

Time: 0830-1230 (Half Day - Morning)



Summary: Computer vision has ceased to be a purely academic endeavor. From law enforcement, to border control, to employment, healthcare diagnostics, and assigning trust scores, computer vision systems have started to be used in all aspects of society. This last year has also seen a rise in public discourse regarding the use of computer-vision based technology by companies such as Google, Microsoft, Amazon and IBM. In research, some works purport to determine a person's sexuality from their social network profile images, and others claim to classify "violent individuals" from drone footage. These works were published in high impact journals, and some were presented at workshops in top tier computer vision conferences such as CVPR.

On the other hand, seminal works published last year showed that commercial gender classification systems have high disparities in error rates by skin-type and gender, exposed the gender bias contained in current image captioning based works, and proposed methods to mitigate bias. Policy makers and other legislators have cited some of these seminal works in their calls to investigate unregulated usage of computer vision systems.

Our workshop aims to provide a space to analyze controversial research papers that have garnered a lot of attention, and highlight research on uncovering and mitigating issues of unfair bias and historical discrimination that trained machine learning models learn to mimic and propagate.

DAVIS Challenge on Video Object Segmentation

Organizers: Sergi Caelles
Jordi Pont-Tuset
Federico Perazzi
Alberto Montes
Kevis-Kokitsi Maninis
Luc Van Gool

Location: Hyatt Beacon B

Time: 0845-1230 (Half Day - Morning)



Summary: We present the 2019 DAVIS Challenge on Video Object Segmentation, a public competition designed for the task of video object segmentation. In addition to the original semi-supervised track and the interactive track introduced in the previous edition, a new unsupervised multi-object track will be featured this year. In the newly introduced track, participants are asked to provide non-overlapping object proposals along with their id for each frame of a video sequence (i.e. video object proposals), without any human supervision. In order to do so, we have re-annotated the training and validation sets of DAVIS 2017 in a concise way that facilitates the unsupervised track, and created new test-dev and test-challenge sets for the competition.

Long-Term Visual Localization Under Changing Conditions

Organizers: Torsten Sattler
Vassileios Balntas
Lars Hammarstrand
Huub Heijnen
Fredrik Kahl
Will Maddern
Krystian Mikolajczyk

Tomas Pajdla
Marc Pollefeys
Johannes L. Schönberger
Josef Sivic
Pablo Speciale
Carl Toft
Akihiko Torii

Location: Hyatt Beacon B

Time: 1330-1800 (Half Day - Afternoon)

Summary: Visual localization is the problem of (accurately) estimating the position and orientation, i.e., the camera pose, from which an image was taken with respect to some scene representation. Visual localization is a vital component in many interesting Computer Vision and Robotics scenarios, including autonomous vehicles such as self-driving cars and other robots, Augmented / Mixed / Virtual Reality, Structure-from-Motion, and SLAM.

Visual localization algorithms rely on a scene representation constructed from images. Since it is impractical to capture a given scene under all potential viewing conditions, i.e., under all potential viewpoints under all potential illumination conditions under all potential seasonal or other conditions, localization algorithms need to be robust to such changes. This is especially true if visual localization algorithms need to operate over a long period of time. This workshop thus focuses on the problem of long-term visual localization and is intended as a benchmark for the current state of visual localization under changing conditions. The workshop consist of both invited talks by experts in the field and practical challenges on recent datasets.



International Skin Imaging Collaboration (ISIC) Workshop on Skin Image Analysis

Organizers: Noel C. F. Codella
M. Emre Celebi
Kristin Dana
Allan Halpern
Philipp Tschandl

Location: Hyatt Seaview B

Time: 1330-1800 (Half Day - Afternoon)

Summary: Skin is the largest organ of the human body, and is the first area of assessment performed by any clinical staff when a patient is seen, as it delivers numerous insights into a patient's underlying health. For example, cardiac function, liver function, immune function, and physical injuries can be assessed by examining the skin. In addition, dermatologic complaints are also among the most prevalent in primary care [1]. Images of the skin are the most easily captured form of medical image in healthcare, and the domain shares qualities to other standard computer vision datasets, serving as a natural bridge between standard computer vision tasks and medical applications.

This workshop will serve as a venue to facilitate advancements and knowledge dissemination in the field of skin image analysis, raising awareness and interest for these socially valuable tasks. Invited speakers include (confirmed) major influencers in computer vision and skin imaging, and authors of accepted papers.



Computational Cameras and Displays

Organizers: Katherine Bouman
Ioannis Gkioulekas
Sanjeev Koppal

Location: Hyatt Regency C

Time: 1300-1800 (Half Day - Afternoon)

Summary: Computational photography has become an increasingly active area of research within the computer vision community. Within the few last years, the amount of research has grown tremendously with dozens of published papers per year in a variety of vision, optics, and graphics venues. A similar trend can be seen in the emerging field of computational displays – spurred by the widespread availability of precise optical and material fabrication technologies, the research community has begun to investigate the joint design of display optics and computational processing. Such displays are not only designed for human observers but also for computer vision applications, providing high-dimensional structured illumination that varies in space, time, angle, and the color spectrum. This workshop is designed to unite the computational camera and display communities in that it considers to what degree concepts from computational cameras can inform the design of emerging computational displays and vice versa, both focused on applications in computer vision.

The CCD workshop series serves as an annual gathering place for researchers and practitioners who design, build, and use computational cameras, displays, and projector-camera systems for a wide variety of uses. The workshop solicits papers, posters, and demo submissions on all topics relating to projector-camera systems.

GigaVision: When Gigapixel Videography Meets Computer Vision

Organizers: Lu Fang
David J. Brady
Ruiping Wang
Shenghua Gao
Yucheng Guo

Location: Hyatt Seaview A

Time: 1300-1700 (Half Day - Afternoon)

Summary: With the development of deep learning theory and technology, the performance of computer vision algorithms including object detection and tracking, face recognition, and 3D reconstruction have made tremendous progress. However, computer vision technology relies on the valid information from the input image and video, and the performance of the algorithm is essentially constrained by the quality of source image/video. Gigapixel videography plays important role in capturing large-scale dynamic scenes for both macro and micro domains. Benefited from the recent progress of gigapixel camera design, the capture of gigapixel-level image/video becomes more and more convenient. However, along with the emergence of gigapixel-level image/video, the corresponding computer vision tasks remain unsolved, due to the extremely high-resolution, large-scale, huge-data that induced by the gigapixel camera. In particular, the understanding of gigapixel video via classical computer vision tasks such as detection, recognition, tracking, segmentation etc. are urgently demanded.



Computer Vision Applications for Mixed Reality Headsets

Organizers: Marc Pollefeys
Federica Bogo
Johannes Schoenberger
Osman Ulusoy

Location: Hyatt Regency F

Time: 1330-1730 (Half Day - Afternoon)

Summary: Mixed reality headsets such as the Microsoft HoloLens are becoming powerful platforms to develop computer vision applications. HoloLens Research Mode enables computer vision research on device by providing access to all raw image sensor streams -- including depth and IR. As Research Mode is now available since May 2018, we are starting to see several interesting demos and applications being developed for HoloLens.

The goal of this workshop is to bring together students and researchers interested in computer vision for mixed reality applications, spanning a broad range of topics: object and activity recognition, hand and user tracking, SLAM, 3D reconstruction, scene understanding, sensor-based localization, navigation and more.

The workshop will provide a venue to share demos and applications, and learn from each other to build or port applications to mixed reality.



When Blockchain Meets Computer Vision & AI

Organizers: Sharathchandra Pankanti
Pramod Viswanath
Karthik Nandakumar
Nalini Ratha

Location: Hyatt Seaview C

Time: 1330-1800 (Half Day - Afternoon)

Summary: Blockchain is a foundational technology that is revolutionizing the way transactions are conceived, executed, and monetized. While the commercial benefits of blockchain infrastructure are imminent, the underlying technological problems need significant attention from researchers. Of specific interest to researchers in computer vision and artificial intelligence (AI) is the tremendous opportunity to make a connection to these emerging infrastructure capabilities and realize how their skills can be leveraged to make an impact by marrying computer vision/AI and blockchain technologies. This marriage can happen in two ways. Firstly, AI technologies can be exploited to address critical gaps in blockchain platforms such as scaling, modeling, and privacy analysis. Secondly, as the world moves towards increasing decentralization of AI and emergence of AI marketplaces, a blockchain-based infrastructure would be essential to create the necessary trust between diverse stakeholders.

As cameras become ubiquitous and compute power becomes pervasively available, the business world is going to consider camera-based analytics as the de facto information channel to improve the integrity of transactions. Although many security, fairness, trust related computer vision and AI topics have been discussed in the past, there has been no attempt to comprehensively bring together the two most exciting technologies of our time – Computer Vision/AI & Blockchain – to leverage their complementary features, both from scientific and industrial communities. This workshop is aimed to bridge this gap.



Photogrammetric Computer Vision

Organizers: Jan-Michael Frahm
Andrea Fusiello
Ronny Hänsch
Alper Yilmaz

Location: 202A

Time: 1330-1745 (Half Day - Afternoon)

Summary: Both photogrammetry and computer vision, refer to the science of acquiring information about the physical world from image data. During the recent decades both fields have matured and converged, which is best illustrated by the widespread usage of the term "Photogrammetric Computer Vision" (PCV). The PCV 2019 workshop aims at providing a forum for collaboration between the computer vision and photogrammetry communities to discuss modern challenges and ideas, propose new and contemporary benchmarks, elaborate on the overlap to machine learning, mathematics and other related areas, and boost the development in the highly challenging and quickly evolving field of photogrammetric computer vision.



Precognition: Seeing Through the Future

Organizers: Khoa Luu
Kris Kitani
Minh Hoai Nguyen
Hien Van Nguyen
Nemanja Djuric
Utsav Prabhu

Location: Hyatt Shoreline A

Time: 1330-1800 (Half Day - Afternoon)

Summary: Vision-based detection and recognition studies have been recently achieving highly accurate performance and were able to bridge the gap between research and real-world applications. Beyond these well-explored detection and recognition capabilities of modern algorithms, vision-based forecasting will likely be one of the next big research topics in the field of computer vision. Vision-based prediction is one of the critical capabilities of humans, and potential success of automatic vision-based forecasting will empower and unlock human-like capabilities in machines and robots.

One important application is in autonomous driving, where vision-based understanding of a traffic scene and prediction of movement of traffic actors is a critical piece of the puzzle. Another area is medical domain, allowing deep understanding and prediction of future medical conditions of patients. However, despite its relevance for real-world applications, visual forecasting or precognition has not been in the focus of new theoretical studies and practical applications as much as detection and recognition problems.

Through organization of this workshop we aim to facilitate further discussion and interest within the research community regarding this nascent topic. This workshop will discuss recent approaches and research trends not only in anticipating human behavior from videos, but also precognition in other visual applications, such as: medical imaging, health-care, human face aging prediction, early even prediction, autonomous driving forecasting, among others.



Tuesday, June 18

0700–1730 Registration (Promenade Atrium & Plaza)**0730–0900 Breakfast** (Pacific Ballroom)**0800–1000 Setup for Poster Session 1-1P** (Exhibit Hall)**0830–0900 Opening Remarks & Paper Awards**
(Terrace Theater; Simulcast to Grand & Promenade Ballrooms)**0900–1015 Oral Session 1-1A: Deep Learning**
(Terrace Theater)

Papers in this session are in Poster Session 1-1P (Posters 1–12)

Chairs: Bharath Hariharan (*Cornell Univ.*)
Subhransu Maji (*Univ. of Massachusetts at Amherst*)Format (5 min. presentation; 3 min. group questions/3 papers)

1. [0900] Finding Task-Relevant Features for Few-Shot Learning by Category Traversal, *Hongyang Li, David Eigen, Samuel Dodge, Matthew Zeiler, Xiaogang Wang*
2. [0905] Edge-Labeling Graph Neural Network for Few-Shot Learning, *Jongmin Kim, Taesup Kim, Sungwoong Kim, Chang D. Yoo*
3. [0910] Generating Classification Weights With GNN Denoising Autoencoders for Few-Shot Learning, *Spyros Gidaris, Nikos Komodakis*
4. [0918] Kervolutional Neural Networks, *Chen Wang, Jianfei Yang, Lihua Xie, Junsong Yuan*
5. [0923] Why ReLU Networks Yield High-Confidence Predictions Far Away From the Training Data and How to Mitigate the Problem, *Matthias Hein, Maksym Andriushchenko, Julian Bitterwolf*
6. [0928] On the Structural Sensitivity of Deep Convolutional Networks to the Directions of Fourier Basis Functions, *Yusuke Tsuzuku, Issei Sato*
7. [0936] Neural Rejuvenation: Improving Deep Network Training by Enhancing Computational Resource Utilization, *Siyuan Qiao, Zhe Lin, Jianming Zhang, Alan L. Yuille*
8. [0941] Hardness-Aware Deep Metric Learning, *Wenzhao Zheng, Zhaodong Chen, Jiwen Lu, Jie Zhou*
9. [0946] Auto-DeepLab: Hierarchical Neural Architecture Search for Semantic Image Segmentation, *Chenxi Liu, Liang-Chieh Chen, Florian Schroff, Hartwig Adam, Wei Hua, Alan L. Yuille, Li Fei-Fei*
10. [0954] Learning Loss for Active Learning, *Donggeun Yoo, In So Kweon*
11. [0959] Striking the Right Balance With Uncertainty, *Salman Khan, Munawar Hayat, Syed Waqas Zamir, Jianbing Shen, Ling Shao*
12. [1004] AutoAugment: Learning Augmentation Strategies From Data, *Ekin D. Cubuk, Barret Zoph, Dandelion Mané, Vijay Vasudevan, Quoc V. Le*

0900–1015 Oral Session 1-1B: 3D Multiview
(Grand Ballroom)

Papers in this session are in Poster Session 1-1P (Posters 72–83)

Chairs: Philippos Mordohai (*Stevens Institute of Technology*)
Hongdong Li (*Australian National Univ.*)Format (5 min. presentation; 3 min. group questions/3 papers)

1. [0900] SDRSAC: Semidefinite-Based Randomized Approach for Robust Point Cloud Registration Without Correspondences, *Huu M. Le, Thanh-Toan Do, Tuan Hoang, Ngai-Man Cheung*
2. [0905] BAD SLAM: Bundle Adjusted Direct RGB-D SLAM, *Thomas Schöps, Torsten Sattler, Marc Pollefeys*
3. [0910] Revealing Scenes by Inverting Structure From Motion Reconstructions, *Francesco Pittaluga, Sanjeev J. Koppal, Sing Bing Kang, Sudipta N. Sinha*
4. [0918] Strand-Accurate Multi-View Hair Capture, *Giljoo Nam, Chenglei Wu, Min H. Kim, Yaser Sheikh*
5. [0923] DeepSDF: Learning Continuous Signed Distance Functions for Shape Representation, *Jeong Joon Park, Peter Florence, Julian Straub, Richard Newcombe, Steven Lovegrove*
6. [0928] Pushing the Boundaries of View Extrapolation With Multiplane Images, *Pratul P. Srinivasan, Richard Tucker, Jonathan T. Barron, Ravi Ramamoorthi, Ren Ng, Noah Snavely*
7. [0936] GA-Net: Guided Aggregation Net for End-To-End Stereo Matching, *Feihu Zhang, Victor Prisacariu, Ruigang Yang, Philip H.S. Torr*
8. [0941] Real-Time Self-Adaptive Deep Stereo, *Alessio Tonioni, Fabio Tosi, Matteo Poggi, Stefano Mattocchia, Luigi Di Stefano*
9. [0946] LAF-Net: Locally Adaptive Fusion Networks for Stereo Confidence Estimation, *Sunok Kim, Seungryong Kim, Dongbo Min, Kwanghoon Sohn*
10. [0954] NM-Net: Mining Reliable Neighbors for Robust Feature Correspondences, *Chen Zhao, Zhiguo Cao, Chi Li, Xin Li, Jiaqi Yang*
11. [0959] Coordinate-Free Carlsson-Weinshall Duality and Relative Multi-View Geometry, *Matthew Trager, Martial Hebert, Jean Ponce*
12. [1004] Deep Reinforcement Learning of Volume-Guided Progressive View Inpainting for 3D Point Scene Completion From a Single Depth Image, *Xiaoguang Han, Zhaoxuan Zhang, Dong Du, Mingdai Yang, Jingming Yu, Pan Pan, Xin Yang, Ligang Liu, Zixiang Xiong, Shuguang Cui*

0900–1015 Oral Session 1-1C: Action & Video
(Promenade Ballroom)

Papers in this session are in Poster Session 1-1P (Posters 109–120)

Chairs: Michael Ryoo (*Google Brain; Indiana Univ.*)
Juan Carlos Niebles (*Stanford Univ.*)Format (5 min. presentation; 3 min. group questions/3 papers)

1. [0900] Video Action Transformer Network, *Rohit Girdhar, João Carreira, Carl Doersch, Andrew Zisserman*
2. [0905] Timeception for Complex Action Recognition, *Noureddien Hussein, Efstratios Gavves, Arnold W.M. Smeulders*
3. [0910] STEP: Spatio-Temporal Progressive Learning for Video Action Detection, *Xitong Yang, Xiaodong Yang, Ming-Yu Liu, Fanyu Xiao, Larry S. Davis, Jan Kautz*
4. [0918] Relational Action Forecasting, *Chen Sun, Abhinav Shrivastava, Carl Vondrick, Rahul Sukthankar, Kevin Murphy, Cordelia Schmid*

5. [0923] Long-Term Feature Banks for Detailed Video Understanding, *Chao-Yuan Wu, Christoph Feichtenhofer, Haoqi Fan, Kaiming He, Philipp Krähenbühl, Ross Girshick*
6. [0928] Which Way Are You Going? Imitative Decision Learning for Path Forecasting in Dynamic Scenes, *Yuke Li*
7. [0936] What and How Well You Performed? A Multitask Learning Approach to Action Quality Assessment, *Paritosh Parmar, Brendan Tran Morris*
8. [0941] MHP-VOS: Multiple Hypotheses Propagation for Video Object Segmentation, *Shuangjie Xu, Daizong Liu, Linchao Bao, Wei Liu, Pan Zhou*
9. [0946] 2.5D Visual Sound, *Ruohan Gao, Kristen Grauman*
10. [0954] Language-Driven Temporal Activity Localization: A Semantic Matching Reinforcement Learning Model, *Weining Wang, Yan Huang, Liang Wang*
11. [0959] Gaussian Temporal Awareness Networks for Action Localization, *Fuchen Long, Ting Yao, Zhaofan Qiu, Xinmei Tian, Jiebo Luo, Tao Mei*
12. [1004] Efficient Video Classification Using Fewer Frames, *Shweta Bhardwaj, Mukundhan Srinivasan, Mitesh M. Khapra*
7. Neural Rejuvenation: Improving Deep Network Training by Enhancing Computational Resource Utilization, *Siyuan Qiao, Zhe Lin, Jianming Zhang, Alan L. Yuille*
8. Hardness-Aware Deep Metric Learning, *Wenzhao Zheng, Zhaodong Chen, Jiwen Lu, Jie Zhou*
9. Auto-DeepLab: Hierarchical Neural Architecture Search for Semantic Image Segmentation, *Chenxi Liu, Liang-Chieh Chen, Florian Schroff, Hartwig Adam, Wei Hua, Alan L. Yuille, Li Fei-Fei*
10. Learning Loss for Active Learning, *Donggeun Yoo, In So Kweon*
11. Striking the Right Balance With Uncertainty, *Salman Khan, Munawar Hayat, Syed Waqas Zamir, Jianbing Shen, Ling Shao*
12. AutoAugment: Learning Augmentation Strategies From Data, *Ekin D. Cubuk, Barret Zoph, Dandelion Mané, Vijay Vasudevan, Quoc V. Le*
13. Parsing R-CNN for Instance-Level Human Analysis, *Lu Yang, Qing Song, Zhihui Wang, Ming Jiang*
14. Large Scale Incremental Learning, *Yue Wu, Yinpeng Chen, Lijuan Wang, Yuancheng Ye, Zicheng Liu, Yandong Guo, Yun Fu*
15. TopNet: Structural Point Cloud Decoder, *Lyne P. Tchapmi, Vineet Kosaraju, Hamid RezaTofighi, Ian Reid, Silvio Savarese*
16. Perceive Where to Focus: Learning Visibility-Aware Part-Level Features for Partial Person Re-Identification, *Yifan Sun, Qin Xu, Yali Li, Chi Zhang, Yikang Li, Shengjin Wang, Jian Sun*
17. Meta-Transfer Learning for Few-Shot Learning, *Qianru Sun, Yaoyao Liu, Tat-Seng Chua, Bernt Schiele*
18. Structured Binary Neural Networks for Accurate Image Classification and Semantic Segmentation, *Bohan Zhuang, Chunhua Shen, Mingkui Tan, Lingqiao Liu, Ian Reid*
19. Deep RNN Framework for Visual Sequential Applications, *Bo Pang, Kaiwen Zha, Hanwen Cao, Chen Shi, Cewu Lu*
20. Graph-Based Global Reasoning Networks, *Yunpeng Chen, Marcus Rohrbach, Zhicheng Yan, Yan Shuicheng, Jiashi Feng, Yannis Kalantidis*
21. SSN: Learning Sparse Switchable Normalization via SparsestMax, *Wenqi Shao, Tianjian Meng, Jingyu Li, Ruimao Zhang, Yudian Li, Xiaogang Wang, Ping Luo*
22. Spherical Fractal Convolutional Neural Networks for Point Cloud Recognition, *Yongming Rao, Jiwen Lu, Jie Zhou*
23. Learning to Generate Synthetic Data via Compositing, *Shashank Tripathi, Siddhartha Chandra, Amit Agrawal, Amrith Tyagi, James M. Rehg, Visesh Chari*
24. Divide and Conquer the Embedding Space for Metric Learning, *Artsiom Sanakoyeu, Vadim Tschernezki, Uta Büchler, Björn Ommer*
25. Latent Space Autoregression for Novelty Detection, *Davide Abati, Angelo Porrello, Simone Calderara, Rita Cucchiara*
26. Attending to Discriminative Certainty for Domain Adaptation, *Vinod Kumar Kurmi, Shanu Kumar, Vinay P. Namboodiri*
27. Feature Denoising for Improving Adversarial Robustness, *Cihang Xie, Yuxin Wu, Laurens van der Maaten, Alan L. Yuille, Kaiming He*
28. Selective Kernel Networks, *Xiang Li, Wenhai Wang, Xiaolin Hu, Jian Yang*
29. On Implicit Filter Level Sparsity in Convolutional Neural Networks, *Dushyant Mehta, Kwang In Kim, Christian Theobalt*
30. FlowNet3D: Learning Scene Flow in 3D Point Clouds, *Xingyu Liu, Charles R. Qi, Leonidas J. Guibas*
31. Scene Memory Transformer for Embodied Agents in Long-Horizon Tasks, *Kuan Fang, Alexander Toshev, Li Fei-Fei, Silvio Savarese*

1015–1115 Morning Break (Exhibit Hall)

1015–1300 Demos (Exhibit Hall)

- Toronto Annotation Suite, *Sanja Fidler, Huan Ling, Jun Gao, Amlan Kar, David Acuna (Univ. of Toronto)*
- A Bio-Inspired Metalens Depth Sensor, *Qi Guo, Zhujun Shi, Yao-Wei Huang, Emma Alexander, Federico Capasso, Todd Zickler (Harvard Univ.)*
- Coded Two-Bucket Cameras for Computer Vision, *Mian Wei, Kiriaikos N. Kutulakos (Univ. of Toronto)*
- Fool The Bank, *Abhishek Bhandwadar, Narendra Nath Joshi, Pin-Yu Chen, Casey Dugan (IBM Research)*

1015–1300 Exhibits (Exhibit Hall)

- See Exhibits map for list of exhibitors.

1015–1300 Poster Session 1-1P (Exhibit Hall)

Deep Learning

1. Finding Task-Relevant Features for Few-Shot Learning by Category Traversal, *Hongyang Li, David Eigen, Samuel Dodge, Matthew Zeiler, Xiaogang Wang*
2. Edge-Labeling Graph Neural Network for Few-Shot Learning, *Jongmin Kim, Taesup Kim, Sungwoong Kim, Chang D. Yoo*
3. Generating Classification Weights With GNN Denoising Autoencoders for Few-Shot Learning, *Spyros Gidaris, Nikos Komodakis*
4. Kervolutional Neural Networks, *Chen Wang, Jianfei Yang, Lihua Xie, Junsong Yuan*
5. Why ReLU Networks Yield High-Confidence Predictions Far Away From the Training Data and How to Mitigate the Problem, *Matthias Hein, Maksym Andriushchenko, Julian Bitterwolf*
6. On the Structural Sensitivity of Deep Convolutional Networks to the Directions of Fourier Basis Functions, *Yusuke Tsuzuku, Issei Sato*

32. Co-Occurrent Features in Semantic Segmentation, *Hang Zhang, Han Zhang, Chenguang Wang, Junyuan Xie*
33. Bag of Tricks for Image Classification with Convolutional Neural Networks, *Tong He, Zhi Zhang, Hang Zhang, Zhongyue Zhang, Junyuan Xie, Mu Li*
34. Learning Channel-Wise Interactions for Binary Convolutional Neural Networks, *Ziwei Wang, Jiwen Lu, Chenxin Tao, Jie Zhou, Qi Tian*
35. Knowledge Adaptation for Efficient Semantic Segmentation, *Tong He, Chunhua Shen, Zhi Tian, Dong Gong, Changming Sun, Youliang Yan*
36. Parametric Noise Injection: Trainable Randomness to Improve Deep Neural Network Robustness Against Adversarial Attack, *Zhezhi He, Adnan Siraj Rakin, Deliang Fan*
- Recognition**
37. Invariance Matters: Exemplar Memory for Domain Adaptive Person Re-Identification, *Zhun Zhong, Liang Zheng, Zhiming Luo, Shaozi Li, Yi Yang*
38. Dissecting Person Re-Identification From the Viewpoint of Viewpoint, *Xiaoxiao Sun, Liang Zheng*
39. Learning to Reduce Dual-Level Discrepancy for Infrared-Visible Person Re-Identification, *Zhixiang Wang, Zheng Wang, Yinqiang Zheng, Yung-Yu Chuang, Shin'ichi Satoh*
40. Progressive Feature Alignment for Unsupervised Domain Adaptation, *Chaoqi Chen, Weiping Xie, Wenbing Huang, Yu Rong, Xinghao Ding, Yue Huang, Tingyang Xu, Junzhou Huang*
41. Feature-Level Frankenstein: Eliminating Variations for Discriminative Recognition, *Xiaofeng Liu, Site Li, Lingsheng Kong, Wanqing Xie, Ping Jia, Jane You, B.V.K. Kumar*
42. Learning a Deep ConvNet for Multi-Label Classification With Partial Labels, *Thibaut Durand, Nazanin Mehrasa, Greg Mori*
43. Generalized Intersection Over Union: A Metric and a Loss for Bounding Box Regression, *Hamid Rezaatofghi, Nathan Tsoi, JunYoung Gwak, Amir Sadeghian, Ian Reid, Silvio Savarese*
44. Densely Semantically Aligned Person Re-Identification, *Zhizheng Zhang, Cuiling Lan, Wenjun Zeng, Zhibo Chen*
45. Generalising Fine-Grained Sketch-Based Image Retrieval, *Kaiyue Pang, Ke Li, Yongxin Yang, Honggang Zhang, Timothy M. Hospedales, Tao Xiang, Yi-Zhe Song*
46. Adapting Object Detectors via Selective Cross-Domain Alignment, *Xinge Zhu, Jiangmiao Pang, Ceyuan Yang, Jianping Shi, Dahua Lin*
47. Cyclic Guidance for Weakly Supervised Joint Detection and Segmentation, *Yunhang Shen, Rongrong Ji, Yan Wang, Yongjian Wu, Liujuan Cao*
48. Thinking Outside the Pool: Active Training Image Creation for Relative Attributes, *Aron Yu, Kristen Grauman*
49. Generalizable Person Re-Identification by Domain-Invariant Mapping Network, *Jifei Song, Yongxin Yang, Yi-Zhe Song, Tao Xiang, Timothy M. Hospedales*
50. Visual Attention Consistency Under Image Transforms for Multi-Label Image Classification, *Hao Guo, Kang Zheng, Xiaochuan Fan, Hongkai Yu, Song Wang*
51. Re-Ranking via Metric Fusion for Object Retrieval and Person Re-Identification, *Song Bai, Peng Tang, Philip H.S. Torr, Longin Jan Latecki*
52. Unsupervised Open Domain Recognition by Semantic Discrepancy Minimization, *Junbao Zhuo, Shuhui Wang, Shuhao Cui, Qingming Huang*
53. Weakly Supervised Person Re-Identification, *Jingke Meng, Sheng Wu, Wei-Shi Zheng*
54. PointRCNN: 3D Object Proposal Generation and Detection From Point Cloud, *Shaoshuai Shi, Xiaogang Wang, Hongsheng Li*
55. Automatic Adaptation of Object Detectors to New Domains Using Self-Training, *Aruni RoyChowdhury, Prithvijit Chakrabarty, Ashish Singh, SouYoung Jin, Huaizu Jiang, Liangliang Cao, Erik Learned-Miller*
56. Deep Sketch-Shape Hashing With Segmented 3D Stochastic Viewing, *Jiaxin Chen, Jie Qin, Li Liu, Fan Zhu, Fumin Shen, Jin Xie, Ling Shao*
57. Generative Dual Adversarial Network for Generalized Zero-Shot Learning, *He Huang, Changhu Wang, Philip S. Yu, Chang-Dong Wang*
58. Query-Guided End-To-End Person Search, *Bharti Munjal, Sikandar Amin, Federico Tombari, Fabio Galasso*
59. Libra R-CNN: Towards Balanced Learning for Object Detection, *Jiangmiao Pang, Kai Chen, Jianping Shi, Huajun Feng, Wanli Ouyang, Dahua Lin*
60. Learning a Unified Classifier Incrementally via Rebalancing, *Saihui Hou, Xinyu Pan, Chen Change Loy, Zilei Wang, Dahua Lin*
61. Feature Selective Anchor-Free Module for Single-Shot Object Detection, *Chenchen Zhu, Yihui He, Marios Savvides*
62. Bottom-Up Object Detection by Grouping Extreme and Center Points, *Xingyi Zhou, Jiacheng Zhuo, Philipp Krähenbühl*
63. Feature Distillation: DNN-Oriented JPEG Compression Against Adversarial Examples, *Zihao Liu, Qi Liu, Tao Liu, Nuo Xu, Xue Lin, Yanzhi Wang, Wujie Wen*
- Segmentation, Grouping, & Shape**
64. SCOPS: Self-Supervised Co-Part Segmentation, *Wei-Chih Hung, Varun Jampani, Sifei Liu, Pavlo Molchanov, Ming-Hsuan Yang, Jan Kautz*
65. Unsupervised Moving Object Detection via Contextual Information Separation, *Yanchao Yang, Antonio Loquercio, Davide Scaramuzza, Stefano Soatto*
66. Pose2Seg: Detection Free Human Instance Segmentation, *Song-Hai Zhang, Ruilong Li, Xin Dong, Paul Rosin, Zixi Cai, Xi Han, Dingcheng Yang, Haozhi Huang, Shi-Min Hu*
- Statistics, Physics, Theory, & Datasets**
67. DrivingStereo: A Large-Scale Dataset for Stereo Matching in Autonomous Driving Scenarios, *Guorun Yang, Xiao Song, Chaoqin Huang, Zhidong Deng, Jianping Shi, Bolei Zhou*
68. PartNet: A Large-Scale Benchmark for Fine-Grained and Hierarchical Part-Level 3D Object Understanding, *Kaichun Mo, Shilin Zhu, Angel X. Chang, Li Yi, Subarna Tripathi, Leonidas J. Guibas, Hao Su*
69. A Dataset and Benchmark for Large-Scale Multi-Modal Face Anti-Spoofing, *Shifeng Zhang, Xiaobo Wang, Ajian Liu, Chenxu Zhao, Jun Wan, Sergio Escalera, Hailin Shi, Zezheng Wang, Stan Z. Li*
70. Unsupervised Learning of Consensus Maximization for 3D Vision Problems, *Thomas Probst, Danda Pani Paudel, Ajad Chhatkuli, Luc Van Gool*
71. VizWiz-Priv: A Dataset for Recognizing the Presence and Purpose of Private Visual Information in Images Taken by Blind People, *Danna Gurari, Qing Li, Chi Lin, Yinan Zhao, Anhong Guo, Abigale Stangl, Jeffrey P. Bigham*

3D Multiview

72. SDRSAC: Semidefinite-Based Randomized Approach for Robust Point Cloud Registration Without Correspondences, *Huu M. Le, Thanh-Toan Do, Tuan Hoang, Ngai-Man Cheung*
73. BAD SLAM: Bundle Adjusted Direct RGB-D SLAM, *Thomas Schöps, Torsten Sattler, Marc Pollefeys*
74. Revealing Scenes by Inverting Structure From Motion Reconstructions, *Francesco Pittaluga, Sanjeev J. Koppal, Sing Bing Kang, Sudipta N. Sinha*
75. Strand-Accurate Multi-View Hair Capture, *Giljoo Nam, Chenglei Wu, Min H. Kim, Yaser Sheikh*
76. DeepSDF: Learning Continuous Signed Distance Functions for Shape Representation, *Jeong Joon Park, Peter Florence, Julian Straub, Richard Newcombe, Steven Lovegrove*
77. Pushing the Boundaries of View Extrapolation With Multiplane Images, *Pratul P. Srinivasan, Richard Tucker, Jonathan T. Barron, Ravi Ramamoorthi, Ren Ng, Noah Snavely*
78. GA-Net: Guided Aggregation Net for End-To-End Stereo Matching, *Feihu Zhang, Victor Prisacariu, Ruigang Yang, Philip H.S. Torr*
79. Real-Time Self-Adaptive Deep Stereo, *Alessio Tonioni, Fabio Tosi, Matteo Poggi, Stefano Mattoccia, Luigi Di Stefano*
80. LAF-Net: Locally Adaptive Fusion Networks for Stereo Confidence Estimation, *Sunok Kim, Seungryong Kim, Dongbo Min, Kwanghoon Sohn*
81. NM-Net: Mining Reliable Neighbors for Robust Feature Correspondences, *Chen Zhao, Zhiguo Cao, Chi Li, Xin Li, Jiaqi Yang*
82. Coordinate-Free Carlsson-Weinshall Duality and Relative Multi-View Geometry, *Matthew Trager, Martial Hebert, Jean Ponce*
83. Deep Reinforcement Learning of Volume-Guided Progressive View Inpainting for 3D Point Scene Completion From a Single Depth Image, *Xiaoguang Han, Zhaoxuan Zhang, Dong Du, Mingdai Yang, Jingming Yu, Pan Pan, Xin Yang, Ligang Liu, Zixiang Xiong, Shuguang Cui*
84. Structural Relational Reasoning of Point Clouds, *Yueqi Duan, Yu Zheng, Jiwen Lu, Jie Zhou, Qi Tian*
85. MVF-Net: Multi-View 3D Face Morphable Model Regression, *Fanzi Wu, Linchao Bao, Yajing Chen, Yonggen Ling, Yibing Song, Songnan Li, King Ng Ngan, Wei Liu*
86. Photometric Mesh Optimization for Video-Aligned 3D Object Reconstruction, *Chen-Hsuan Lin, Oliver Wang, Bryan C. Russell, Eli Shechtman, Vladimir G. Kim, Matthew Fisher, Simon Lucey*
87. Guided Stereo Matching, *Matteo Poggi, Davide Pallotti, Fabio Tosi, Stefano Mattoccia*
88. Unsupervised Event-Based Learning of Optical Flow, Depth, and Egomotion, *Alex Zihao Zhu, Liangzhe Yuan, Kenneth Chaney, Kostas Daniilidis*
89. Modeling Local Geometric Structure of 3D Point Clouds Using Geo-CNN, *Shiyi Lan, Ruichi Yu, Gang Yu, Larry S. Davis*

3D Single View & RGBD

90. 3D Point Capsule Networks, *Yongheng Zhao, Tolga Birdal, Haowen Deng, Federico Tombari*
91. GS3D: An Efficient 3D Object Detection Framework for Autonomous Driving, *Buyu Li, Wanli Ouyang, Lu Sheng, Xingyu Zeng, Xiaogang Wang*
92. Single-Image Piece-Wise Planar 3D Reconstruction via Associative Embedding, *Zehao Yu, Jia Zheng, Dongze Lian, Zihan Zhou, Shenghua Gao*

93. 3DN: 3D Deformation Network, *Weiyue Wang, Duygu Ceylan, Radomir Mech, Ulrich Neumann*
94. HorizonNet: Learning Room Layout With 1D Representation and Pano Stretch Data Augmentation, *Cheng Sun, Chi-Wei Hsiao, Min Sun, Hwann-Tzong Chen*
95. Deep Fitting Degree Scoring Network for Monocular 3D Object Detection, *Lijie Liu, Jiwen Lu, Chunjing Xu, Qi Tian, Jie Zhou*

Face & Body

96. Pushing the Envelope for RGB-Based Dense 3D Hand Pose Estimation via Neural Rendering, *Seungryul Baek, Kwang In Kim, Tae-Kyun Kim*
97. Self-Supervised Learning of 3D Human Pose Using Multi-View Geometry, *Muhammed Kocabas, Salih Karagoz, Emre Akbas*
98. FSA-Net: Learning Fine-Grained Structure Aggregation for Head Pose Estimation From a Single Image, *Tsun-Yi Yang, Yi-Ting Chen, Yen-Yu Lin, Yung-Yu Chuang*
99. Dense 3D Face Decoding Over 2500FPS: Joint Texture & Shape Convolutional Mesh Decoders, *Yuxiang Zhou, Jiankang Deng, Irene Kotsia, Stefanos Zafeiriou*
100. Does Learning Specific Features for Related Parts Help Human Pose Estimation? *Wei Tang, Ying Wu*
101. Linkage Based Face Clustering via Graph Convolution Network, *Zhongdao Wang, Liang Zheng, Yali Li, Shengjin Wang*
102. Towards High-Fidelity Nonlinear 3D Face Morphable Model, *Luan Tran, Feng Liu, Xiaoming Liu*
103. RegularFace: Deep Face Recognition via Exclusive Regularization, *Kai Zhao, Jingyi Xu, Ming-Ming Cheng*
104. BridgeNet: A Continuity-Aware Probabilistic Network for Age Estimation, *Wanhua Li, Jiwen Lu, Jianjiang Feng, Chunjing Xu, Jie Zhou, Qi Tian*
105. GANFIT: Generative Adversarial Network Fitting for High Fidelity 3D Face Reconstruction, *Baris Gecer, Stylianos Ploumpis, Irene Kotsia, Stefanos Zafeiriou*
106. Improving the Performance of Unimodal Dynamic Hand-Gesture Recognition With Multimodal Training, *Mahdi Abavisani, Hamid Reza Vaezi Joze, Vishal M. Patel*
107. Learning to Reconstruct People in Clothing From a Single RGB Camera, *Thiemo Alldieck, Marcus Magnor, Bharat Lal Bhatnagar, Christian Theobalt, Gerard Pons-Moll*
108. Distilled Person Re-Identification: Towards a More Scalable System, *Ancong Wu, Wei-Shi Zheng, Xiaowei Guo, Jian-Huang Lai*

Action & Video

109. Video Action Transformer Network, *Rohit Girdhar, João Carreira, Carl Doersch, Andrew Zisserman*
110. Timeception for Complex Action Recognition, *Noureddien Hussein, Efstratios Gavves, Arnold W.M. Smeulders*
111. STEP: Spatio-Temporal Progressive Learning for Video Action Detection, *Xitong Yang, Xiaodong Yang, Ming-Yu Liu, Fanyi Xiao, Larry S. Davis, Jan Kautz*
112. Relational Action Forecasting, *Chen Sun, Abhinav Shrivastava, Carl Vondrick, Rahul Sukthankar, Kevin Murphy, Cordelia Schmid*
113. Long-Term Feature Banks for Detailed Video Understanding, *Chao-Yuan Wu, Christoph Feichtenhofer, Haoqi Fan, Kaiming He, Philipp Krähenbühl, Ross Girshick*
114. Which Way Are You Going? Imitative Decision Learning for Path Forecasting in Dynamic Scenes, *Yuke Li*
115. What and How Well You Performed? A Multitask Learning Approach to Action Quality Assessment, *Paritosh Parmar, Brendan Tran Morris*

116. MHP-VOS: Multiple Hypotheses Propagation for Video Object Segmentation, *Shuangjie Xu, Daizong Liu, Linchao Bao, Wei Liu, Pan Zhou*
117. 2.5D Visual Sound, *Ruohan Gao, Kristen Grauman*
118. Language-Driven Temporal Activity Localization: A Semantic Matching Reinforcement Learning Model, *Weining Wang, Yan Huang, Liang Wang*
119. Gaussian Temporal Awareness Networks for Action Localization, *Fuchen Long, Ting Yao, Zhaofan Qiu, Xinmei Tian, Jiebo Luo, Tao Mei*
120. Efficient Video Classification Using Fewer Frames, *Shweta Bhardwaj, Mukundhan Srinivasan, Mitesh M. Khapra*
121. A Perceptual Prediction Framework for Self Supervised Event Segmentation, *Sathyannarayanan N. Aakur, Sudeep Sarkar*
122. COIN: A Large-Scale Dataset for Comprehensive Instructional Video Analysis, *Yansong Tang, Dajun Ding, Yongming Rao, Yu Zheng, Danyang Zhang, Lili Zhao, Jiwen Lu, Jie Zhou*
123. Recurrent Attentive Zooming for Joint Crowd Counting and Precise Localization, *Chenchen Liu, Xinyu Weng, Yadong Mu*
124. An Attention Enhanced Graph Convolutional LSTM Network for Skeleton-Based Action Recognition, *Chenyang Si, Wentao Chen, Wei Wang, Liang Wang, Tieniu Tan*
125. Graph Convolutional Label Noise Cleaner: Train a Plug-And-Play Action Classifier for Anomaly Detection, *Jia-Xing Zhong, Nannan Li, Weijie Kong, Shan Liu, Thomas H. Li, Ge Li*
126. MAN: Moment Alignment Network for Natural Language Moment Retrieval via Iterative Graph Adjustment, *Da Zhang, Xiyang Dai, Xin Wang, Yuan-Fang Wang, Larry S. Davis*
127. Less Is More: Learning Highlight Detection From Video Duration, *Bo Xiong, Yannis Kalantidis, Deepti Ghadiyaram, Kristen Grauman*
128. DMC-Net: Generating Discriminative Motion Cues for Fast Compressed Video Action Recognition, *Zheng Shou, Xudong Lin, Yannis Kalantidis, Laura Sevilla-Lara, Marcus Rohrbach, Shih-Fu Chang, Zhicheng Yan*
129. AdaFrame: Adaptive Frame Selection for Fast Video Recognition, *Zuxuan Wu, Caiming Xiong, Chih-Yao Ma, Richard Socher, Larry S. Davis*
130. Spatio-Temporal Video Re-Localization by Warp LSTM, *Yang Feng, Lin Ma, Wei Liu, Jiebo Luo*
131. Completeness Modeling and Context Separation for Weakly Supervised Temporal Action Localization, *Daochang Liu, Tingting Jiang, Yizhou Wang*

Motion & Biometrics

132. Unsupervised Deep Tracking, *Ning Wang, Yibing Song, Chao Ma, Wengang Zhou, Wei Liu, Houqiang Li*
133. Tracking by Animation: Unsupervised Learning of Multi-Object Attentive Trackers, *Zhen He, Jian Li, Daxue Liu, Hangen He, David Barber*
134. Fast Online Object Tracking and Segmentation: A Unifying Approach, *Qiang Wang, Li Zhang, Luca Bertinetto, Weiming Hu, Philip H.S. Torr*
135. Object Tracking by Reconstruction With View-Specific Discriminative Correlation Filters, *Uğur Kart, Alan Lukežič, Matej Kristan, Joni-Kristian Kämäräinen, Jiří Matas*
136. SoPhie: An Attentive GAN for Predicting Paths Compliant to Social and Physical Constraints, *Amir Sadeghian, Vineet Kosaraju, Ali Sadeghian, Noriaki Hirose, Hamid Rezaatofighi, Silvio Savarese*

137. Leveraging Shape Completion for 3D Siamese Tracking, *Silvio Giancola, Jesus Zarzar, Bernard Ghanem*
138. Target-Aware Deep Tracking, *Xin Li, Chao Ma, Baoyuan Wu, Zhenyu He, Ming-Hsuan Yang*
139. Spatiotemporal CNN for Video Object Segmentation, *Kai Xu, Longyin Wen, Guorong Li, Liefeng Bo, Qingming Huang*
140. Towards Rich Feature Discovery With Class Activation Maps Augmentation for Person Re-Identification, *Wenjie Yang, Houjing Huang, Zhang Zhang, Xiaotang Chen, Kaiqi Huang, Shu Zhang*

Synthesis

141. Wide-Context Semantic Image Extrapolation, *Yi Wang, Xin Tao, Xiaoyong Shen, Jiaya Jia*
142. End-To-End Time-Lapse Video Synthesis From a Single Outdoor Image, *Seonghyeon Nam, Chongyang Ma, Menglei Chai, William Brendel, Ning Xu, Seon Joo Kim*
143. GIF2Video: Color Dequantization and Temporal Interpolation of GIF Images, *Yang Wang, Haibin Huang, Chuan Wang, Tong He, Jue Wang, Minh Hoai*
144. Mode Seeking Generative Adversarial Networks for Diverse Image Synthesis, *Qi Mao, Hsin-Ying Lee, Hung-Yu Tseng, Siwei Ma, Ming-Hsuan Yang*
145. Pluralistic Image Completion, *Chuanxia Zheng, Tat-Jen Cham, Jianfei Cai*
146. Salient Object Detection With Pyramid Attention and Salient Edges, *Wenguan Wang, Shuyang Zhao, Jianbing Shen, Steven C. H. Hoi, Ali Borji*
147. Latent Filter Scaling for Multimodal Unsupervised Image-To-Image Translation, *Yazeed Alharbi, Neil Smith, Peter Wonka*
148. Attention-Aware Multi-Stroke Style Transfer, *Yuan Yao, Jianqiang Ren, Xuansong Xie, Weidong Liu, Yong-Jin Liu, Jun Wang*
149. Feedback Adversarial Learning: Spatial Feedback for Improving Generative Adversarial Networks, *Minyoung Huh, Shao-Hua Sun, Ning Zhang*
150. Learning Pyramid-Context Encoder Network for High-Quality Image Inpainting, *Yanhong Zeng, Jianlong Fu, Hongyang Chao, Baining Guo*
151. Example-Guided Style-Consistent Image Synthesis From Semantic Labeling, *Miao Wang, Guo-Ye Yang, Ruilong Li, Run-Ze Liang, Song-Hai Zhang, Peter M. Hall, Shi-Min Hu*
152. MirrorGAN: Learning Text-To-Image Generation by Redescription, *Tingting Qiao, Jing Zhang, Duanqing Xu, Dacheng Tao*

Computational Photography & Graphics

153. Light Field Messaging With Deep Photographic Steganography, *Eric Wengrowski, Kristin Dana*
154. Im2Pencil: Controllable Pencil Illustration From Photographs, *Yijun Li, Chen Fang, Aaron Hertzmann, Eli Shechtman, Ming-Hsuan Yang*
155. When Color Constancy Goes Wrong: Correcting Improperly White-Balanced Images, *Mahmoud Afffi, Brian Price, Scott Cohen, Michael S. Brown*
156. Beyond Volumetric Albedo — A Surface Optimization Framework for Non-Line-Of-Sight Imaging, *Chia-Yin Tsai, Aswin C. Sankaranarayanan, Ioannis Gkioulekas*
157. Reflection Removal Using a Dual-Pixel Sensor, *Abhijith Punnappurath, Michael S. Brown*

158. Practical Coding Function Design for Time-Of-Flight Imaging, *Felipe Gutierrez-Barragan, Syed Azer Reza, Andreas Velten, Mohit Gupta*
159. Meta-SR: A Magnification-Arbitrary Network for Super-Resolution, *Xuecai Hu, Haoyuan Mu, Xiangyu Zhang, Zilei Wang, Tieniu Tan, Jian Sun*

Low-Level & Optimization

160. Multispectral and Hyperspectral Image Fusion by MS/HS Fusion Net, *Qi Xie, Minghao Zhou, Qian Zhao, Deyu Meng, Wangmeng Zuo, Zongben Xu*
161. Learning Attraction Field Representation for Robust Line Segment Detection, *Nan Xue, Song Bai, Fudong Wang, Gui-Song Xia, Tianfu Wu, Liangpei Zhang*
162. Blind Super-Resolution With Iterative Kernel Correction, *Jinjin Gu, Hannan Lu, Wangmeng Zuo, Chao Dong*
163. Video Magnification in the Wild Using Fractional Anisotropy in Temporal Distribution, *Shoichiro Takeda, Yasunori Akagi, Kazuki Okami, Megumi Isogai, Hideaki Kimata*
164. Attentive Feedback Network for Boundary-Aware Salient Object Detection, *Mengyang Feng, Huchuan Lu, Errui Ding*
165. Heavy Rain Image Restoration: Integrating Physics Model and Conditional Adversarial Learning, *Ruoteng Li, Loong-Fah Cheong, Robby T. Tan*
166. Learning to Calibrate Straight Lines for Fisheye Image Rectification, *Zhucun Xue, Nan Xue, Gui-Song Xia, Weiming Shen*
167. Camera Lens Super-Resolution, *Chang Chen, Zhiwei Xiong, Xinmei Tian, Zheng-Jun Zha, Feng Wu*
168. Frame-Consistent Recurrent Video Deraining With Dual-Level Flow, *Wenhan Yang, Jiaying Liu, Jiashi Feng*
169. Deep Plug-And-Play Super-Resolution for Arbitrary Blur Kernels, *Kai Zhang, Wangmeng Zuo, Lei Zhang*
170. Sea-Thru: A Method for Removing Water From Underwater Images, *Derya Akkaynak, Tali Treibitz*
171. Deep Network Interpolation for Continuous Imagery Effect Transition, *Xintao Wang, Ke Yu, Chao Dong, Xiaoou Tang, Chen Change Loy*
172. Spatially Variant Linear Representation Models for Joint Filtering, *Jinshan Pan, Jiangxin Dong, Jimmy S. Ren, Liang Lin, Jinhui Tang, Ming-Hsuan Yang*
173. Toward Convolutional Blind Denoising of Real Photographs, *Shi Guo, Zifei Yan, Kai Zhang, Wangmeng Zuo, Lei Zhang*
174. Towards Real Scene Super-Resolution With Raw Images, *Xiangyu Xu, Yongrui Ma, Wenxiu Sun*
175. ODE-Inspired Network Design for Single Image Super-Resolution, *Xiangyu He, Zitao Mo, Peisong Wang, Yang Liu, Mingyuan Yang, Jian Cheng*
176. Blind Image Deblurring With Local Maximum Gradient Prior, *Liang Chen, Faming Fang, Tingting Wang, Guixu Zhang*
177. Attention-Guided Network for Ghost-Free High Dynamic Range Imaging, *Qingsen Yan, Dong Gong, Qinfeng Shi, Anton van den Hengel, Chunhua Shen, Ian Reid, Yanning Zhang*
181. CrDoCo: Pixel-Level Domain Transfer With Cross-Domain Consistency, *Yun-Chun Chen, Yen-Yu Lin, Ming-Hsuan Yang, Jia-Bin Huang*
182. Temporal Cycle-Consistency Learning, *Debidatta Dwibedi, Yusuf Aytar, Jonathan Tompson, Pierre Sermanet, Andrew Zisserman*
183. Predicting Future Frames Using Retrospective Cycle GAN, *Yong-Hoon Kwon, Min-Gyu Park*
184. Density Map Regression Guided Detection Network for RGB-D Crowd Counting and Localization, *Dongze Lian, Jing Li, Jia Zheng, Weixin Luo, Shenghua Gao*
185. TAFE-Net: Task-Aware Feature Embeddings for Low Shot Learning, *Xin Wang, Fisher Yu, Ruth Wang, Trevor Darrell, Joseph E. Gonzalez*
186. Learning Semantic Segmentation From Synthetic Data: A Geometrically Guided Input-Output Adaptation Approach, *Yuhua Chen, Wen Li, Xiaoran Chen, Luc Van Gool*
187. Attentive Single-Tasking of Multiple Tasks, *Kevis-Kokitsi Maninis, Ilija Radosavovic, Iasonas Kokkinos*
188. Deep Metric Learning to Rank, *Fatih Cakir, Kun He, Xide Xia, Brian Kulis, Stan Sclaroff*
189. End-To-End Multi-Task Learning With Attention, *Shikun Liu, Edward Johns, Andrew J. Davison*
190. Self-Supervised Learning via Conditional Motion Propagation, *Xiaohang Zhan, Xingang Pan, Ziwei Liu, Dahua Lin, Chen Change Loy*
191. Bridging Stereo Matching and Optical Flow via Spatiotemporal Correspondence, *Hsueh-Ying Lai, Yi-Hsuan Tsai, Wei-Chen Chiu*
192. All About Structure: Adapting Structural Information Across Domains for Boosting Semantic Segmentation, *Wei-Lun Chang, Hui-Po Wang, Wen-Hsiao Peng, Wei-Chen Chiu*
193. Iterative Reorganization With Weak Spatial Constraints: Solving Arbitrary Jigsaw Puzzles for Unsupervised Representation Learning, *Chen Wei, Lingxi Xie, Xutong Ren, Yingda Xia, Chi Su, Jiaying Liu, Qi Tian, Alan L. Yuille*
194. Revisiting Self-Supervised Visual Representation Learning, *Alexander Kolesnikov, Xiaohua Zhai, Lucas Beyer*

Language & Reasoning

Scenes & Representation

178. Searching for a Robust Neural Architecture in Four GPU Hours, *Xuanyi Dong, Yi Yang*
179. Hierarchy Denoising Recursive Autoencoders for 3D Scene Layout Prediction, *Yifei Shi, Angel X. Chang, Zhelun Wu, Manolis Savva, Kai Xu*
180. Adaptively Connected Neural Networks, *Guangrun Wang, Keze Wang, Liang Lin*

195. It's Not About the Journey; It's About the Destination: Following Soft Paths Under Question-Guidance for Visual Reasoning, *Monica Haurilet, Alina Roitberg, Rainer Stiefelhagen*
196. Actively Seeking and Learning From Live Data, *Damien Teney, Anton van den Hengel*
197. Improving Referring Expression Grounding With Cross-Modal Attention-Guided Erasing, *Xihui Liu, Zihao Wang, Jing Shao, Xiaogang Wang, Hongsheng Li*
198. Neighbourhood Watch: Referring Expression Comprehension via Language-Guided Graph Attention Networks, *Peng Wang, Qi Wu, Jiewei Cao, Chunhua Shen, Lianli Gao, Anton van den Hengel*
199. Scene Graph Generation With External Knowledge and Image Reconstruction, *Jiuxiang Gu, Handong Zhao, Zhe Lin, Sheng Li, Jianfei Cai, Mingyang Ling*
200. Polysemous Visual-Semantic Embedding for Cross-Modal Retrieval, *Yale Song, Mohammad Soleymani*
201. MUREL: Multimodal Relational Reasoning for Visual Question Answering, *Remi Cadene, Hedi Ben-younes, Matthieu Cord, Nicolas Thome*
202. Heterogeneous Memory Enhanced Multimodal Attention Model for Video Question Answering, *Chenyou Fan, Xiaofan Zhang, Shu Zhang, Wensheng Wang, Chi Zhang, Heng Huang*

1320–1520 Setup for Poster Session 1-2P (Exhibit Hall)**1330–1520 Oral Session 1-2A: Recognition**

(Terrace Theater)

Papers in this session are in Poster Session 1-2P (Posters 18–35)

Chairs: Zeynep Akata (*Univ. of Amsterdam*)
Jia Deng (*Princeton Univ.*)**Format** (5 min. presentation; 3 min. group questions/3 papers)

1. [1330] Joint Discriminative and Generative Learning for Person Re-Identification, *Zhedong Zheng, Xiaodong Yang, Zhiding Yu, Liang Zheng, Yi Yang, Jan Kautz*
2. [1335] Unsupervised Person Re-Identification by Soft Multilabel Learning, *Hong-Xing Yu, Wei-Shi Zheng, Ancong Wu, Xiaowei Guo, Shaogang Gong, Jian-Huang Lai*
3. [1340] Learning Context Graph for Person Search, *Yichao Yan, Qiang Zhang, Bingbing Ni, Wendong Zhang, Minghao Xu, Xiaokang Yang*
4. [1348] Gradient Matching Generative Networks for Zero-Shot Learning, *Mert Bulent Sariyildiz, Ramazan Gokberk Cinbis*
5. [1353] Doodle to Search: Practical Zero-Shot Sketch-Based Image Retrieval, *Sounak Dey, Pau Riba, Anjan Dutta, Josep Lladós, Yi-Zhe Song*
6. [1358] Zero-Shot Task Transfer, *Arghya Pal, Vineeth N Balasubramanian*
7. [1406] C-MIL: Continuation Multiple Instance Learning for Weakly Supervised Object Detection, *Fang Wan, Chang Liu, Wei Ke, Xiangyang Ji, Jianbin Jiao, Qixiang Ye*
8. [1411] Weakly Supervised Learning of Instance Segmentation With Inter-Pixel Relations, *Jiwoon Ahn, Sunghyun Cho, Suha Kwak*
9. [1416] Attention-Based Dropout Layer for Weakly Supervised Object Localization, *Junsuk Choe, Hyunjung Shim*
10. [1424] Domain Generalization by Solving Jigsaw Puzzles, *Fabio M. Carlucci, Antonio D'Innocente, Silvia Bucci, Barbara Caputo, Tatiana Tommasi*
11. [1429] Transferrable Prototypical Networks for Unsupervised Domain Adaptation, *Yingwei Pan, Ting Yao, Yehao Li, Yu Wang, Chong-Wah Ngo, Tao Mei*
12. [1434] Blending-Target Domain Adaptation by Adversarial Meta-Adaptation Networks, *Ziliang Chen, Jingyu Zhuang, Xiaodan Liang, Liang Lin*
13. [1442] ELASTIC: Improving CNNs With Dynamic Scaling Policies, *Huiyu Wang, Aniruddha Kembhavi, Ali Farhadi, Alan L. Yuille, Mohammad Rastegari*
14. [1447] ScratchDet: Training Single-Shot Object Detectors From Scratch, *Rui Zhu, Shifeng Zhang, Xiaobo Wang, Longyin Wen, Hailin Shi, Liefeng Bo, Tao Mei*
15. [1452] SFNet: Learning Object-Aware Semantic Correspondence, *Junghyup Lee, Dohyung Kim, Jean Ponce, Bumsub Ham*
16. [1500] Deep Metric Learning Beyond Binary Supervision, *Sungyeon Kim, Minkyoo Seo, Ivan Laptev, Minsu Cho, Suha Kwak*
17. [1505] Learning to Cluster Faces on an Affinity Graph, *Lei Yang, Xiaohang Zhan, Dapeng Chen, Junjie Yan, Chen Change Loy, Dahua Lin*
18. [1510] C2AE: Class Conditioned Auto-Encoder for Open-Set Recognition, *Poojan Oza, Vishal M. Patel*

1330–1520 Oral Session 1-2B: Synthesis

(Grand Ballroom)

Papers in this session are in Poster Session 1-2P (Posters 118–135)

Chairs: Philip Isola (*Massachusetts Institute of Technology*)
James Hays (*Georgia Institute of Technology*)**Format** (5 min. presentation; 3 min. group questions/3 papers)

1. [1330] Shapes and Context: In-The-Wild Image Synthesis & Manipulation, *Aayush Bansal, Yaser Sheikh, Deva Ramanan*
2. [1335] Semantics Disentangling for Text-To-Image Generation, *Guojun Yin, Bin Liu, Lu Sheng, Nenghai Yu, Xiaogang Wang, Jing Shao*
3. [1340] Semantic Image Synthesis With Spatially-Adaptive Normalization, *Taesung Park, Ming-Yu Liu, Ting-Chun Wang, Jun-Yan Zhu*
4. [1348] Progressive Pose Attention Transfer for Person Image Generation, *Zhen Zhu, Tengting Huang, Baoguang Shi, Miao Yu, Bofei Wang, Xiang Bai*
5. [1353] Unsupervised Person Image Generation With Semantic Parsing Transformation, *Sijie Song, Wei Zhang, Jiaying Liu, Tao Mei*
6. [1358] DeepView: View Synthesis With Learned Gradient Descent, *John Flynn, Michael Broxton, Paul Debevec, Matthew DuVall, Graham Fyffe, Ryan Overbeck, Noah Snively, Richard Tucker*
7. [1406] Animating Arbitrary Objects via Deep Motion Transfer, *Aliaksandr Siarohin, Stéphane Lathuilière, Sergey Tulyakov, Elisa Ricci, Nicu Sebe*
8. [1411] Textured Neural Avatars, *Aliaksandra Shysheya, Egor Zakharov, Kara-Ali Aliev, Renat Bashirov, Egor Burkov, Karim Isakov, Aleksei Ivakhnenko, Yury Malkov, Igor Pasechnik, Dmitry Ulyanov, Alexander Vakhitov, Victor Lempitsky*
9. [1416] IM-Net for High Resolution Video Frame Interpolation, *Tomer Peleg, Pablo Szekely, Doron Sabo, Omry Sendik*
10. [1424] Homomorphic Latent Space Interpolation for Unpaired Image-To-Image Translation, *Ying-Cong Chen, Xiaogang Xu, Zhuotao Tian, Jiaya Jia*
11. [1429] Multi-Channel Attention Selection GAN With Cascaded Semantic Guidance for Cross-View Image Translation, *Hao Tang, Dan Xu, Nicu Sebe, Yanzhi Wang, Jason J. Corso, Yan Yan*
12. [1434] Geometry-Consistent Generative Adversarial Networks for One-Sided Unsupervised Domain Mapping, *Huan Fu, Mingming Gong, Chaohui Wang, Kayhan Batmanghelich, Kun Zhang, Dacheng Tao*
13. [1442] DeepVoxels: Learning Persistent 3D Feature Embeddings, *Vincent Sitzmann, Justus Thies, Felix Heide, Matthias Nießner, Gordon Wetzstein, Michael Zollhöfer*
14. [1447] Inverse Path Tracing for Joint Material and Lighting Estimation, *Dejan Azinović, Tzu-Mao Li, Anton Kaplanyan, Matthias Nießner*
15. [1452] The Visual Centrifuge: Model-Free Layered Video Representations, *Jean-Baptiste Alayrac, João Carreira, Andrew Zisserman*
16. [1500] Label-Noise Robust Generative Adversarial Networks, *Takuhiko Kaneko, Yoshitaka Ushiku, Tatsuya Harada*
17. [1505] DLOW: Domain Flow for Adaptation and Generalization, *Rui Gong, Wen Li, Yuhua Chen, Luc Van Gool*
18. [1510] CollaGAN: Collaborative GAN for Missing Image Data Imputation, *Dongwook Lee, Junyoung Kim, Won-Jin Moon, Jong Chul Ye*

1330–1520 Oral Session 1-2C: Scenes & Representation (Promenade Ballroom)

Papers in this session are in Poster Session 1-2P (Posters 166–183)

Chairs: Qixing Huang (*Univ. of Texas at Austin*)
Hao Su (*Univ. of California, San Diego*)

Format (5 min. presentation; 3 min. group questions/3 papers)

1. [1330] d-SNE: Domain Adaptation Using Stochastic Neighborhood Embedding, *Xiang Xu, Xiong Zhou, Ragav Venkatesan, Gurumurthy Swaminathan, Orchid Majumder*
2. [1335] Taking a Closer Look at Domain Shift: Category-Level Adversaries for Semantics Consistent Domain Adaptation, *Yawei Luo, Liang Zheng, Tao Guan, Junqing Yu, Yi Yang*
3. [1340] ADVENT: Adversarial Entropy Minimization for Domain Adaptation in Semantic Segmentation, *Tuan-Hung Vu, Himalaya Jain, Maxime Bucher, Matthieu Cord, Patrick Pérez*

4. [1348] ContextDesc: Local Descriptor Augmentation With Cross-Modality Context, *Zixin Luo, Tianwei Shen, Lei Zhou, Jiahui Zhang, Yao Yao, Shiwei Li, Tian Fang, Long Quan*
5. [1353] Large-Scale Long-Tailed Recognition in an Open World, *Ziwei Liu, Zhongqi Miao, Xiaohang Zhan, Jiayun Wang, Boqing Gong, Stella X. Yu*
6. [1358] AET vs. AED: Unsupervised Representation Learning by Auto-Encoding Transformations Rather Than Data, *Liheng Zhang, Guo-Jun Qi, Liqiang Wang, Jiebo Luo*

7. [1406] SDC – Stacked Dilated Convolution: A Unified Descriptor Network for Dense Matching Tasks, *René Schuster, Oliver Wasenmüller, Christian Unger, Didier Stricker*
8. [1411] Learning Correspondence From the Cycle-Consistency of Time, *Xiaolong Wang, Allan Jabri, Alexei A. Efros*
9. [1416] AE2-Nets: Autoencoder in Autoencoder Networks, *Changqing Zhang, Yeqing Liu, Huazhu Fu*

10. [1424] Mitigating Information Leakage in Image Representations: A Maximum Entropy Approach, *Proteek Chandan Roy, Vishnu Naresh Boddeti*
11. [1429] Learning Spatial Common Sense With Geometry-Aware Recurrent Networks, *Hsiao-Yu Fish Tung, Ricson Cheng, Katerina Fragkiadaki*
12. [1434] Structured Knowledge Distillation for Semantic Segmentation, *Yifan Liu, Ke Chen, Chris Liu, Zengchang Qin, Zhenbo Luo, Jingdong Wang*

13. [1442] Scan2CAD: Learning CAD Model Alignment in RGB-D Scans, *Armen Avetisyan, Manuel Dahnert, Angela Dai, Manolis Savva, Angel X. Chang, Matthias Nießner*
14. [1447] Towards Scene Understanding: Unsupervised Monocular Depth Estimation With Semantic-Aware Representation, *Po-Yi Chen, Alexander H. Liu, Yen-Cheng Liu, Yu-Chiang Frank Wang*
15. [1452] Tell Me Where I Am: Object-Level Scene Context Prediction, *Xiaotian Qiao, Quanlong Zheng, Ying Cao, Rynson W.H. Lau*

16. [1500] Normalized Object Coordinate Space for Category-Level 6D Object Pose and Size Estimation, *He Wang, Srinath Sridhar, Jingwei Huang, Julien Valentin, Shuran Song, Leonidas J. Guibas*
17. [1505] Supervised Fitting of Geometric Primitives to 3D Point Clouds, *Lingxiao Li, Minhyuk Sung, Anastasia Dubrovina, Li Yi, Leonidas J. Guibas*
18. [1510] Do Better ImageNet Models Transfer Better? *Simon Kornblith, Jonathon Shlens, Quoc V. Le*

1520–1620 Afternoon Break (Exhibit Hall)

1520–1800 Demos (Exhibit Hall)

- Local Light Field Fusion, *Ben Mildenhall, Rodrigo Ortiz-Cayon, Abhishek Kar (Fyusion Inc, UC Berkeley)*
- 3D-Printed Object Recognition, *Sebastian Koch (Technische Universität Berlin)*
- Interactive Fashion Retrieval Using Natural Language Feedback, *Xiaoxiao Guo, Yupeng Gao, Hui Wu (IBM Research)*
- Explanation-Assisted Guess Which, *Arijit Ray, Giedrius Burachas, Yi Yao, Ajay Divakaran (SRI International)*

1520–1800 Exhibits (Exhibit Hall)

- See Exhibits map for list of exhibitors.

1520–1800 Poster Session 1-2P (Exhibit Hall)

Deep Learning

1. Gotta Adapt 'Em All: Joint Pixel and Feature-Level Domain Adaptation for Recognition in the Wild, *Luan Tran, Kihyuk Sohn, Xiang Yu, Xiaoming Liu, Manmohan Chandraker*
2. Understanding the Disharmony Between Dropout and Batch Normalization by Variance Shift, *Xiang Li, Shuo Chen, Xiaolin Hu, Jian Yang*
3. Circulant Binary Convolutional Networks: Enhancing the Performance of 1-Bit DCNNs With Circulant Back Propagation, *Chunlei Liu, Wenrui Ding, Xin Xia, Baochang Zhang, Jiabin Gu, Jianzhuang Liu, Rongrong Ji, David Doermann*
4. DeFusionNET: Defocus Blur Detection via Recurrently Fusing and Refining Multi-Scale Deep Features, *Chang Tang, Xinzhong Zhu, Xinwang Liu, Lizhe Wang, Albert Zomaya*
5. Deep Virtual Networks for Memory Efficient Inference of Multiple Tasks, *Eunwoo Kim, Chanho Ahn, Philip H.S. Torr, Songhwai Oh*
6. Universal Domain Adaptation, *Kaichao You, Mingsheng Long, Zhangjie Cao, Jianmin Wang, Michael I. Jordan*
7. Improving Transferability of Adversarial Examples With Input Diversity, *Cihang Xie, Zhishuai Zhang, Yuyin Zhou, Song Bai, Jianyu Wang, Zhou Ren, Alan L. Yuille*
8. Sequence-To-Sequence Domain Adaptation Network for Robust Text Image Recognition, *Yaping Zhang, Shuai Nie, Wenju Liu, Xing Xu, Dongxiang Zhang, Heng Tao Shen*
9. Hybrid-Attention Based Decoupled Metric Learning for Zero-Shot Image Retrieval, *Binghui Chen, Weihong Deng*
10. Learning to Sample, *Oren Dovrat, Itai Lang, Shai Avidan*
11. Few-Shot Learning via Saliency-Guided Hallucination of Samples, *Hongguang Zhang, Jing Zhang, Piotr Koniusz*
12. Variational Convolutional Neural Network Pruning, *Chenglong Zhao, Bingbing Ni, Jian Zhang, Qiwei Zhao, Wenjun Zhang, Qi Tian*
13. Towards Optimal Structured CNN Pruning via Generative Adversarial Learning, *Shaohui Lin, Rongrong Ji, Chenqian Yan, Baochang Zhang, Liujuan Cao, Qixiang Ye, Feiyue Huang, David Doermann*
14. Exploiting Kernel Sparsity and Entropy for Interpretable CNN Compression, *Yuchao Li, Shaohui Lin, Baochang Zhang, Jianzhuang Liu, David Doermann, Yongjian Wu, Feiyue Huang, Rongrong Ji*

15. Fully Quantized Network for Object Detection, *Rundong Li, Yan Wang, Feng Liang, Hongwei Qin, Junjie Yan, Rui Fan*
 16. MnasNet: Platform-Aware Neural Architecture Search for Mobile, *Mingxing Tan, Bo Chen, Ruoming Pang, Vijay Vasudevan, Mark Sandler, Andrew Howard, Quoc V. Le*
 17. Student Becoming the Master: Knowledge Amalgamation for Joint Scene Parsing, Depth Estimation, and More, *Jingwen Ye, Yixin Ji, Xinchao Wang, Kairi Ou, Dapeng Tao, Mingli Song*
- Recognition**
18. Joint Discriminative and Generative Learning for Person Re-Identification, *Zhedong Zheng, Xiaodong Yang, Zhiding Yu, Liang Zheng, Yi Yang, Jan Kautz*
 19. Unsupervised Person Re-Identification by Soft Multilabel Learning, *Hong-Xing Yu, Wei-Shi Zheng, Ancong Wu, Xiaowei Guo, Shaogang Gong, Jian-Huang Lai*
 20. Learning Context Graph for Person Search, *Yichao Yan, Qiang Zhang, Bingbing Ni, Wendong Zhang, Minghao Xu, Xiaokang Yang*
 21. Gradient Matching Generative Networks for Zero-Shot Learning, *Mert Bulent Sariyildiz, Ramazan Gokberk Cinbis*
 22. Doodle to Search: Practical Zero-Shot Sketch-Based Image Retrieval, *Sounak Dey, Pau Riba, Anjan Dutta, Josep Lladós, Yi-Zhe Song*
 23. Zero-Shot Task Transfer, *Arghya Pal, Vineeth N Balasubramanian*
 24. C-MIL: Continuation Multiple Instance Learning for Weakly Supervised Object Detection, *Fang Wan, Chang Liu, Wei Ke, Xiangyang Ji, Jianbin Jiao, Qixiang Ye*
 25. Weakly Supervised Learning of Instance Segmentation With Inter-Pixel Relations, *Jiwoon Ahn, Sunghyun Cho, Suha Kwak*
 26. Attention-Based Dropout Layer for Weakly Supervised Object Localization, *Junsuk Choe, Hyunjung Shim*
 27. Domain Generalization by Solving Jigsaw Puzzles, *Fabio M. Carlucci, Antonio D'Innocente, Silvia Bucci, Barbara Caputo, Tatiana Tommasi*
 28. Transferrable Prototypical Networks for Unsupervised Domain Adaptation, *Yingwei Pan, Ting Yao, Yehao Li, Yu Wang, Chong-Wah Ngo, Tao Mei*
 29. Blending-Target Domain Adaptation by Adversarial Meta-Adaptation Networks, *Ziliang Chen, Jingyu Zhuang, Xiaodan Liang, Liang Lin*
 30. ELASTIC: Improving CNNs With Dynamic Scaling Policies, *Huiyu Wang, Aniruddha Kembhavi, Ali Farhadi, Alan L. Yuille, Mohammad Rastegari*
 31. ScratchDet: Training Single-Shot Object Detectors From Scratch, *Rui Zhu, Shifeng Zhang, Xiaobo Wang, Longyin Wen, Hailin Shi, Liefeng Bo, Tao Mei*
 32. SFNet: Learning Object-Aware Semantic Correspondence, *Junghyup Lee, Dohyung Kim, Jean Ponce, Bumsuh Ham*
 33. Deep Metric Learning Beyond Binary Supervision, *Sungyeon Kim, Minkyoo Seo, Ivan Laptev, Minsu Cho, Suha Kwak*
 34. Learning to Cluster Faces on an Affinity Graph, *Lei Yang, Xiaohang Zhan, Dapeng Chen, Junjie Yan, Chen Change Loy, Dahua Lin*
 35. C2AE: Class Conditioned Auto-Encoder for Open-Set Recognition, *Poojan Oza, Vishal M. Patel*
 36. K-Nearest Neighbors Hashing, *Xiangyu He, Peisong Wang, Jian Cheng*
 37. Learning RoI Transformer for Oriented Object Detection in Aerial Images, *Jian Ding, Nan Xue, Yang Long, Gui-Song Xia, Qikai Lu*
 38. Snapshot Distillation: Teacher-Student Optimization in One Generation, *Chenglin Yang, Lingxi Xie, Chi Su, Alan L. Yuille*
 39. Geometry-Aware Distillation for Indoor Semantic Segmentation, *Jianbo Jiao, Yunchao Wei, Zequn Jie, Honghui Shi, Rynson W.H. Lau, Thomas S. Huang*
 40. LiveSketch: Query Perturbations for Guided Sketch-Based Visual Search, *John Collomosse, Tu Bui, Hailin Jin*
 41. Bounding Box Regression With Uncertainty for Accurate Object Detection, *Yihui He, Chenchen Zhu, Jianren Wang, Marios Savvides, Xiangyu Zhang*
 42. OCGAN: One-Class Novelty Detection Using GANs With Constrained Latent Representations, *Pramuditha Perera, Ramesh Nallapati, Bing Xiang*
 43. Learning Metrics From Teachers: Compact Networks for Image Embedding, *Lu Yu, Vacit Oguz Yazici, Xialei Liu, Joost van de Weijer, Yongmei Cheng, Arnau Ramisa*
 44. Activity Driven Weakly Supervised Object Detection, *Zhenheng Yang, Dhruv Mahajan, Deepti Ghadiyaram, Ram Nevatia, Vignesh Ramanathan*
 45. Separate to Adapt: Open Set Domain Adaptation via Progressive Separation, *Hong Liu, Zhangjie Cao, Mingsheng Long, Jianmin Wang, Qiang Yang*
 46. Layout-Graph Reasoning for Fashion Landmark Detection, *Weijiang Yu, Xiaodan Liang, Ke Gong, Chenhan Jiang, Nong Xiao, Liang Lin*
 47. DistillHash: Unsupervised Deep Hashing by Distilling Data Pairs, *Erkun Yang, Tongliang Liu, Cheng Deng, Wei Liu, Dacheng Tao*
 48. Mind Your Neighbours: Image Annotation With Metadata Neighbourhood Graph Co-Attention Networks, *Junjie Zhang, Qi Wu, Jian Zhang, Chunhua Shen, Jianfeng Lu*
 49. Region Proposal by Guided Anchoring, *Jiaqi Wang, Kai Chen, Shuo Yang, Chen Change Loy, Dahua Lin*
 50. Distant Supervised Centroid Shift: A Simple and Efficient Approach to Visual Domain Adaptation, *Jian Liang, Ran He, Zhenan Sun, Tieniu Tan*
 51. Learning to Transfer Examples for Partial Domain Adaptation, *Zhangjie Cao, Kaichao You, Mingsheng Long, Jianmin Wang, Qiang Yang*
 52. Generalized Zero-Shot Recognition Based on Visually Semantic Embedding, *Pengkai Zhu, Hanxiao Wang, Venkatesh Saligrama*
 53. Towards Visual Feature Translation, *Jie Hu, Rongrong Ji, Hong Liu, Shengchuan Zhang, Cheng Deng, Qi Tian*
 54. Amodal Instance Segmentation With KINS Dataset, *Lu Qi, Li Jiang, Shu Liu, Xiaoyong Shen, Jiaya Jia*
 55. Global Second-Order Pooling Convolutional Networks, *Zilin Gao, Jiangtao Xie, Qilong Wang, Peihua Li*
 56. Weakly Supervised Complementary Parts Models for Fine-Grained Image Classification From the Bottom Up, *Weifeng Ge, Xiangru Lin, Yizhou Yu*
 57. NetTailor: Tuning the Architecture, Not Just the Weights, *Pedro Morgado, Nuno Vasconcelos*
- Segmentation, Grouping, & Shape**
58. Learning-Based Sampling for Natural Image Matting, *Jingwei Tang, Yağiz Aksoy, Cengiz Öztireli, Markus Gross, Tunç Ozan Aydin*

59. Learning Unsupervised Video Object Segmentation Through Visual Attention, *Wenguan Wang, Hongmei Song, Shuyang Zhao, Jianbing Shen, Sanyuan Zhao, Steven C. H. Hoi, Haibin Ling*
60. 4D Spatio-Temporal ConvNets: Minkowski Convolutional Neural Networks, *Christopher Choy, JunYoung Gwak, Silvio Savarese*
61. Pyramid Feature Attention Network for Saliency Detection, *Ting Zhao, Xiangqian Wu*
62. Co-Saliency Detection via Mask-Guided Fully Convolutional Networks With Multi-Scale Label Smoothing, *Kaihua Zhang, Tengpeng Li, Bo Liu, Qingshan Liu*
63. SAIL-VOS: Semantic Amodal Instance Level Video Object Segmentation – A Synthetic Dataset and Baselines, *Yuan-Ting Hu, Hong-Shuo Chen, Kexin Hui, Jia-Bin Huang, Alexander G. Schwing*
64. Learning Instance Activation Maps for Weakly Supervised Instance Segmentation, *Yi Zhu, Yanzhao Zhou, Huijuan Xu, Qixiang Ye, David Doermann, Jianbin Jiao*
65. Decoders Matter for Semantic Segmentation: Data-Dependent Decoding Enables Flexible Feature Aggregation, *Zhi Tian, Tong He, Chunhua Shen, Youliang Yan*
66. Box-Driven Class-Wise Region Masking and Filling Rate Guided Loss for Weakly Supervised Semantic Segmentation, *Chunfeng Song, Yan Huang, Wanli Ouyang, Liang Wang*
67. Dual Attention Network for Scene Segmentation, *Jun Fu, Jing Liu, Haijie Tian, Yong Li, Yongjun Bao, Zhiwei Fang, Hanqing Lu*

Statistics, Physics, Theory, & Datasets

68. InverseRenderNet: Learning Single Image Inverse Rendering, *Ye Yu, William A. P. Smith*
69. A Variational Auto-Encoder Model for Stochastic Point Processes, *Nazanin Mehrasa, Akash Abdu Jyothi, Thibaut Durand, Jiawei He, Leonid Sigal, Greg Mori*
70. Unifying Heterogeneous Classifiers With Distillation, *Jayakorn Vongkulbhisal, Phongtharin Vinayavekkin, Marco Visentini-Scarzanella*
71. Assessment of Faster R-CNN in Man-Machine Collaborative Search, *Arturo Deza, Amit Surana, Miguel P. Eckstein*
72. OK-VQA: A Visual Question Answering Benchmark Requiring External Knowledge, *Kenneth Marino, Mohammad Rastegari, Ali Farhadi, Roozbeh Mottaghi*
73. NDDR-CNN: Layerwise Feature Fusing in Multi-Task CNNs by Neural Discriminative Dimensionality Reduction, *Yuan Gao, Jiayi Ma, Mingbo Zhao, Wei Liu, Alan L. Yuille*
74. Spectral Metric for Dataset Complexity Assessment, *Frédéric Branchaud-Charron, Andrew Achkar, Pierre-Marc Jodoin*
75. ADCrowdNet: An Attention-Injective Deformable Convolutional Network for Crowd Understanding, *Ning Liu, Yongchao Long, Changqing Zou, Qun Niu, Li Pan, Hefeng Wu*
76. VERI-Wild: A Large Dataset and a New Method for Vehicle Re-Identification in the Wild, *Yihang Lou, Yan Bai, Jun Liu, Shiqi Wang, Lingyu Duan*

3D Multiview

77. 3D Local Features for Direct Pairwise Registration, *Haowen Deng, Tolga Birdal, Slobodan Ilic*
78. HPLFlowNet: Hierarchical Permutohedral Lattice FlowNet for Scene Flow Estimation on Large-Scale Point Clouds, *Xiuye Gu, Yijie Wang, Chongruo Wu, Yong Jae Lee, Panqu Wang*

79. GPSfM: Global Projective SFM Using Algebraic Constraints on Multi-View Fundamental Matrices, *Yoni Kasten, Amnon Geifman, Meirav Galun, Ronen Basri*
80. Group-Wise Correlation Stereo Network, *Xiaoyang Guo, Kai Yang, Wukui Yang, Xiaogang Wang, Hongsheng Li*
81. Multi-Level Context Ultra-Aggregation for Stereo Matching, *Guang-Yu Nie, Ming-Ming Cheng, Yun Liu, Zhengfa Liang, Deng-Ping Fan, Yue Liu, Yongtian Wang*
82. Large-Scale, Metric Structure From Motion for Unordered Light Fields, *Sotiris Nousias, Manolis Lourakis, Christos Bergeles*
83. Understanding the Limitations of CNN-Based Absolute Camera Pose Regression, *Torsten Sattler, Qunjie Zhou, Marc Pollefeys, Laura Leal-Taixé*
84. DeepLiDAR: Deep Surface Normal Guided Depth Prediction for Outdoor Scene From Sparse LiDAR Data and Single Color Image, *Jiaxiong Qiu, Zhaopeng Cui, Yinda Zhang, Xingdi Zhang, Shuaicheng Liu, Bing Zeng, Marc Pollefeys*
85. Modeling Point Clouds With Self-Attention and Gumbel Subset Sampling, *Jiancheng Yang, Qiang Zhang, Bingbing Ni, Linguo Li, Jinxian Liu, Mengdie Zhou, Qi Tian*
86. Learning With Batch-Wise Optimal Transport Loss for 3D Shape Recognition, *Lin Xu, Han Sun, Yuai Liu*
87. DenseFusion: 6D Object Pose Estimation by Iterative Dense Fusion, *Chen Wang, Danfei Xu, Yuke Zhu, Roberto Martín-Martín, Cewu Lu, Li Fei-Fei, Silvio Savarese*

3D Single View & RGBD

88. Dense Depth Posterior (DDP) From Single Image and Sparse Range, *Yanchao Yang, Alex Wong, Stefano Soatto*
89. DuLa-Net: A Dual-Projection Network for Estimating Room Layouts From a Single RGB Panorama, *Shang-Ta Yang, Fu-En Wang, Chi-Han Peng, Peter Wonka, Min Sun, Hung-Kuo Chu*
90. Veritatem Dies Aperit - Temporally Consistent Depth Prediction Enabled by a Multi-Task Geometric and Semantic Scene Understanding Approach, *Amir Atapour-Abarghouei, Toby P. Breckon*
91. Segmentation-Driven 6D Object Pose Estimation, *Yinlin Hu, Joachim Hugonot, Pascal Fua, Mathieu Salzmann*
92. Exploiting Temporal Context for 3D Human Pose Estimation in the Wild, *Anurag Arnab, Carl Doersch, Andrew Zisserman*
93. What Do Single-View 3D Reconstruction Networks Learn? *Maxim Tatarchenko, Stephan R. Richter, René Ranftl, Zhuwen Li, Vladlen Koltun, Thomas Brox*

Face & Body

94. UniformFace: Learning Deep Equidistributed Representation for Face Recognition, *Yueqi Duan, Jiwen Lu, Jie Zhou*
95. Semantic Graph Convolutional Networks for 3D Human Pose Regression, *Long Zhao, Xi Peng, Yu Tian, Mubbasir Kapadia, Dimitris N. Metaxas*
96. Mask-Guided Portrait Editing With Conditional GANs, *Shuyang Gu, Jianmin Bao, Hao Yang, Dong Chen, Fang Wen, Lu Yuan*
97. Group Sampling for Scale Invariant Face Detection, *Xiang Ming, Fangyun Wei, Ting Zhang, Dong Chen, Fang Wen*
98. Joint Representation and Estimator Learning for Facial Action Unit Intensity Estimation, *Yong Zhang, Baoyuan Wu, Weiming Dong, Zhifeng Li, Wei Liu, Bao-Gang Hu, Qiang Ji*
99. Semantic Alignment: Finding Semantically Consistent Ground-Truth for Facial Landmark Detection, *Zhiwei Liu, Xiangyu Zhu, Guosheng Hu, Haiyun Guo, Ming Tang, Zhen Lei, Neil M. Robertson, Jinqiao Wang*

100. LAEO-Net: Revisiting People Looking at Each Other in Videos, *Manuel J. Marín-Jiménez, Vicky Kalogeiton, Pablo Medina-Suárez, Andrew Zisserman*
101. Robust Facial Landmark Detection via Occlusion-Adaptive Deep Networks, *Meilu Zhu, Daming Shi, Mingjie Zheng, Muhammad Sadiq*
102. Learning Individual Styles of Conversational Gesture, *Shiry Ginosar, Amir Bar, Gefen Kohavi, Caroline Chan, Andrew Owens, Jitendra Malik*
103. Face Anti-Spoofing: Model Matters, so Does Data, *Xiao Yang, Wenhan Luo, Linchao Bao, Yuan Gao, Dihong Gong, Shibao Zheng, Zhifeng Li, Wei Liu*
104. Fast Human Pose Estimation, *Feng Zhang, Xiatian Zhu, Mao Ye*
105. Decorrelated Adversarial Learning for Age-Invariant Face Recognition, *Hao Wang, Dihong Gong, Zhifeng Li, Wei Liu*
- Action & Video**
106. Cross-Task Weakly Supervised Learning From Instructional Videos, *Dimitri Zhukov, Jean-Baptiste Alayrac, Ramazan Gokberk Cinbis, David Fouhey, Ivan Laptev, Josef Sivic*
107. D3TW: Discriminative Differentiable Dynamic Time Warping for Weakly Supervised Action Alignment and Segmentation, *Chien-Yi Chang, De-An Huang, Yanan Sui, Li Fei-Fei, Juan Carlos Niebles*
108. Progressive Teacher-Student Learning for Early Action Prediction, *Xionghui Wang, Jian-Fang Hu, Jian-Huang Lai, Jianguo Zhang, Wei-Shi Zheng*
109. Social Relation Recognition From Videos via Multi-Scale Spatial-Temporal Reasoning, *Xinchen Liu, Wu Liu, Meng Zhang, Jingwen Chen, Lianli Gao, Chenggang Yan, Tao Mei*
110. MS-TCN: Multi-Stage Temporal Convolutional Network for Action Segmentation, *Yazan Abu Farha, Jürgen Gall*
111. Transferable Interactiveness Knowledge for Human-Object Interaction Detection, *Yong-Lu Li, Siyuan Zhou, Xijie Huang, Liang Xu, Ze Ma, Hao-Shu Fang, Yanfeng Wang, Cewu Lu*
112. Actional-Structural Graph Convolutional Networks for Skeleton-Based Action Recognition, *Maosen Li, Siheng Chen, Xu Chen, Ya Zhang, Yanfeng Wang, Qi Tian*
113. Multi-Granularity Generator for Temporal Action Proposal, *Yuan Liu, Lin Ma, Yifeng Zhang, Wei Liu, Shih-Fu Chang*
- Motion & Biometrics**
114. Deep Rigid Instance Scene Flow, *Wei-Chiu Ma, Shenlong Wang, Rui Hu, Yuwen Xiong, Raquel Urtasun*
115. See More, Know More: Unsupervised Video Object Segmentation With Co-Attention Siamese Networks, *Xiankai Lu, Wenguan Wang, Chao Ma, Jianbing Shen, Ling Shao, Fatih Porikli*
116. Patch-Based Discriminative Feature Learning for Unsupervised Person Re-Identification, *Qize Yang, Hong-Xing Yu, Ancong Wu, Wei-Shi Zheng*
117. SPM-Tracker: Series-Parallel Matching for Real-Time Visual Object Tracking, *Guangting Wang, Chong Luo, Zhiwei Xiong, Wenjun Zeng*
- Synthesis**
118. Shapes and Context: In-The-Wild Image Synthesis & Manipulation, *Aayush Bansal, Yaser Sheikh, Deva Ramanan*
119. Semantics Disentangling for Text-To-Image Generation, *Guojun Yin, Bin Liu, Lu Sheng, Nenghai Yu, Xiaogang Wang, Jing Shao*
120. Semantic Image Synthesis With Spatially-Adaptive Normalization, *Taesung Park, Ming-Yu Liu, Ting-Chun Wang, Jun-Yan Zhu*
121. Progressive Pose Attention Transfer for Person Image Generation, *Zhen Zhu, Tengpeng Huang, Baoguang Shi, Miao Yu, Bofei Wang, Xiang Bai*
122. Unsupervised Person Image Generation With Semantic Parsing Transformation, *Sijie Song, Wei Zhang, Jiaying Liu, Tao Mei*
123. DeepView: View Synthesis With Learned Gradient Descent, *John Flynn, Michael Broxton, Paul Debevec, Matthew DuVall, Graham Fyffe, Ryan Overbeck, Noah Snavely, Richard Tucker*
124. Animating Arbitrary Objects via Deep Motion Transfer, *Aliaksandr Siarohin, Stéphane Lathuilière, Sergey Tulyakov, Elisa Ricci, Nicu Sebe*
125. Textured Neural Avatars, *Aliaksandra Shysheya, Egor Zakharov, Kara-Ali Aliev, Renat Bashirov, Egor Burkov, Karim Iskakov, Aleksei Ivakhnenko, Yury Malkov, Igor Pasechnik, Dmitry Ulyanov, Alexander Vakhitov, Victor Lempitsky*
126. IM-Net for High Resolution Video Frame Interpolation, *Tomer Peleg, Pablo Szekely, Doron Sabo, Omry Sendik*
127. Homomorphic Latent Space Interpolation for Unpaired Image-To-Image Translation, *Ying-Cong Chen, Xiaogang Xu, Zhuotao Tian, Jiaya Jia*
128. Multi-Channel Attention Selection GAN With Cascaded Semantic Guidance for Cross-View Image Translation, *Hao Tang, Dan Xu, Nicu Sebe, Yanzhi Wang, Jason J. Corso, Yan Yan*
129. Geometry-Consistent Generative Adversarial Networks for One-Sided Unsupervised Domain Mapping, *Huan Fu, Mingming Gong, Chaohui Wang, Kayhan Batmanghelich, Kun Zhang, Dacheng Tao*
130. DeepVoxels: Learning Persistent 3D Feature Embeddings, *Vincent Sitzmann, Justus Thies, Felix Heide, Matthias Nießner, Gordon Wetzstein, Michael Zollhöfer*
131. Inverse Path Tracing for Joint Material and Lighting Estimation, *Dejan Azinović, Tzu-Mao Li, Anton Kaplanyan, Matthias Nießner*
132. The Visual Centrifuge: Model-Free Layered Video Representations, *Jean-Baptiste Alayrac, João Carreira, Andrew Zisserman*
133. Label-Noise Robust Generative Adversarial Networks, *Takuhiro Kaneko, Yoshitaka Ushiku, Tatsuya Harada*
134. DLOW: Domain Flow for Adaptation and Generalization, *Rui Gong, Wen Li, Yuhua Chen, Luc Van Gool*
135. CollaGAN: Collaborative GAN for Missing Image Data Imputation, *Dongwook Lee, Junyoung Kim, Won-Jin Moon, Jong Chul Ye*
136. Spatial Fusion GAN for Image Synthesis, *Fangneng Zhan, Hongyuan Zhu, Shijian Lu*
137. Text Guided Person Image Synthesis, *Xingran Zhou, Siyu Huang, Bin Li, Yingming Li, Jiachen Li, Zhongfei Zhang*
138. STGAN: A Unified Selective Transfer Network for Arbitrary Image Attribute Editing, *Ming Liu, Yukang Ding, Min Xia, Xiao Liu, Errui Ding, Wangmeng Zuo, Shilei Wen*
139. Towards Instance-Level Image-To-Image Translation, *Zhiqiang Shen, Mingyang Huang, Jianping Shi, Xiangyang Xue, Thomas S. Huang*
140. Dense Intrinsic Appearance Flow for Human Pose Transfer, *Yining Li, Chen Huang, Chen Change Loy*

141. Depth-Aware Video Frame Interpolation, *Wenbo Bao, Wei-Sheng Lai, Chao Ma, Xiaoyun Zhang, Zhiyong Gao, Ming-Hsuan Yang*
142. Sliced Wasserstein Generative Models, *Jiqing Wu, Zhiwu Huang, Dinesh Acharya, Wen Li, Janine Thoma, Danda Pani Paudel, Luc Van Gool*
143. Deep Flow-Guided Video Inpainting, *Rui Xu, Xiaoxiao Li, Bolei Zhou, Chen Change Loy*
144. Video Generation From Single Semantic Label Map, *Junting Pan, Chengyu Wang, Xu Jia, Jing Shao, Lu Sheng, Junjie Yan, Xiaogang Wang*

Computational Photography & Graphics

145. Polarimetric Camera Calibration Using an LCD Monitor, *Zhixiang Wang, Yinqiang Zheng, Yung-Yu Chuang*
146. Fully Automatic Video Colorization With Self-Regularization and Diversity, *Chenyang Lei, Qifeng Chen*
147. Zoom to Learn, Learn to Zoom, *Xuaner Zhang, Qifeng Chen, Ren Ng, Vladlen Koltun*
148. Single Image Reflection Removal Beyond Linearity, *Qiang Wen, Yinjie Tan, Jing Qin, Wenxi Liu, Guoqiang Han, Shengfeng He*
149. Learning to Separate Multiple Illuminants in a Single Image, *Zhuo Hui, Ayan Chakrabarti, Kalyan Sunkavalli, Aswin C. Sankaranarayanan*
150. Shape Unicode: A Unified Shape Representation, *Sanjeev Muralikrishnan, Vladimir G. Kim, Matthew Fisher, Siddhartha Chaudhuri*
151. Robust Video Stabilization by Optimization in CNN Weight Space, *Jiyang Yu, Ravi Ramamoorthi*

Low-Level & Optimization

152. Learning Linear Transformations for Fast Image and Video Style Transfer, *Xueting Li, Sifei Liu, Jan Kautz, Ming-Hsuan Yang*
153. Local Detection of Stereo Occlusion Boundaries, *Jialiang Wang, Todd Zickler*
154. Bi-Directional Cascade Network for Perceptual Edge Detection, *Jianzhong He, Shiliang Zhang, Ming Yang, Yanhu Shan, Tiejun Huang*
155. Single Image Deraining: A Comprehensive Benchmark Analysis, *Siyuan Li, Iago Breno Araujo, Wenqi Ren, Zhangyang Wang, Eric K. Tokuda, Roberto Hirata Junior, Roberto Cesar-Junior, Jiawan Zhang, Xiaojie Guo, Xiaochun Cao*
156. Dynamic Scene Deblurring With Parameter Selective Sharing and Nested Skip Connections, *Hongyun Gao, Xin Tao, Xiaoyong Shen, Jiaya Jia*
157. Events-To-Video: Bringing Modern Computer Vision to Event Cameras, *Henri Rebecq, René Ranftl, Vladlen Koltun, Davide Scaramuzza*
158. Feedback Network for Image Super-Resolution, *Zhen Li, Jinglei Yang, Zheng Liu, Xiaomin Yang, Gwanggil Jeon, Wei Wu*
159. Semi-Supervised Transfer Learning for Image Rain Removal, *Wei Wei, Deyu Meng, Qian Zhao, Zongben Xu, Ying Wu*
160. EventNet: Asynchronous Recursive Event Processing, *Yusuke Sekikawa, Kosuke Hara, Hideo Saito*
161. Recurrent Back-Projection Network for Video Super-Resolution, *Muhammad Haris, Gregory Shakhnarovich, Norimichi Ukita*
162. Cascaded Partial Decoder for Fast and Accurate Salient Object Detection, *Zhe Wu, Li Su, Qingming Huang*
163. A Simple Pooling-Based Design for Real-Time Salient Object Detection, *Jiang-Jiang Liu, Qibin Hou, Ming-Ming Cheng, Jiashi Feng, Jianmin Jiang*

164. Contrast Prior and Fluid Pyramid Integration for RGBD Salient Object Detection, *Jia-Xing Zhao, Yang Cao, Deng-Ping Fan, Ming-Ming Cheng, Xuan-Yi Li, Le Zhang*
165. Progressive Image Deraining Networks: A Better and Simpler Baseline, *Dongwei Ren, Wangmeng Zuo, Qinghua Hu, Pengfei Zhu, Deyu Meng*

Scenes & Representation

166. d-SNE: Domain Adaptation Using Stochastic Neighborhood Embedding, *Xiang Xu, Xiong Zhou, Ragav Venkatesan, Gurumurthy Swaminathan, Orchid Majumder*
167. Taking a Closer Look at Domain Shift: Category-Level Adversaries for Semantics Consistent Domain Adaptation, *Yawei Luo, Liang Zheng, Tao Guan, Junqing Yu, Yi Yang*
168. ADVENT: Adversarial Entropy Minimization for Domain Adaptation in Semantic Segmentation, *Tuan-Hung Vu, Himalaya Jain, Maxime Bucher, Matthieu Cord, Patrick Pérez*
169. ContextDesc: Local Descriptor Augmentation With Cross-Modality Context, *Zixin Luo, Tianwei Shen, Lei Zhou, Jiahui Zhang, Yao Yao, Shiwei Li, Tian Fang, Long Quan*
170. Large-Scale Long-Tailed Recognition in an Open World, *Ziwei Liu, Zhongqi Miao, Xiaohang Zhan, Jiayun Wang, Boqing Gong, Stella X. Yu*
171. AET vs. AED: Unsupervised Representation Learning by Auto-Encoding Transformations Rather Than Data, *Liheng Zhang, Guo-Jun Qi, Liqiang Wang, Jiebo Luo*
172. SDC – Stacked Dilated Convolution: A Unified Descriptor Network for Dense Matching Tasks, *René Schuster, Oliver Wasenmüller, Christian Unger, Didier Stricker*
173. Learning Correspondence From the Cycle-Consistency of Time, *Xiaolong Wang, Allan Jabri, Alexei A. Efros*
174. AE2-Nets: Autoencoder in Autoencoder Networks, *Changqing Zhang, Yeqing Liu, Huazhu Fu*
175. Mitigating Information Leakage in Image Representations: A Maximum Entropy Approach, *Proteek Chandan Roy, Vishnu Naresh Boddeti*
176. Learning Spatial Common Sense With Geometry-Aware Recurrent Networks, *Hsiao-Yu Fish Tung, Ricson Cheng, Katerina Fragkiadaki*
177. Structured Knowledge Distillation for Semantic Segmentation, *Yifan Liu, Ke Chen, Chris Liu, Zengchang Qin, Zhenbo Luo, Jingdong Wang*
178. Scan2CAD: Learning CAD Model Alignment in RGB-D Scans, *Armen Avetisyan, Manuel Dahnert, Angela Dai, Manolis Savva, Angel X. Chang, Matthias Nießner*
179. Towards Scene Understanding: Unsupervised Monocular Depth Estimation With Semantic-Aware Representation, *Po-Yi Chen, Alexander H. Liu, Yen-Cheng Liu, Yu-Chiang Frank Wang*
180. Tell Me Where I Am: Object-Level Scene Context Prediction, *Xiaotian Qiao, Quanlong Zheng, Ying Cao, Rynson W.H. Lau*
181. Normalized Object Coordinate Space for Category-Level 6D Object Pose and Size Estimation, *He Wang, Srinath Sridhar, Jingwei Huang, Julien Valentin, Shuran Song, Leonidas J. Guibas*
182. Supervised Fitting of Geometric Primitives to 3D Point Clouds, *Lingxiao Li, Minhyuk Sung, Anastasia Dubrovina, Li Yi, Leonidas J. Guibas*
183. Do Better ImageNet Models Transfer Better? *Simon Kornblith, Jonathon Shlens, Quoc V. Le*

Wednesday, June 19

0730-1730 Registration (Promenade Atrium & Plaza)

0730-0900 Breakfast (Pacific Ballroom)

0800-1000 Setup for Poster Session 2-1P (Exhibit Hall)

0830-1000 Oral Session 2-1A: Deep Learning
(Terrace Theater)

Papers in this session are in Poster Session 2-1P (Posters 1-15)

Chairs: Laurens van der Maaten (*Facebook*)
Zhe Lin (*Adobe Research*)

Format (5 min. presentation; 3 min. group questions/3 papers)

1. [0830] Learning Video Representations From Correspondence Proposals, *Xingyu Liu, Joon-Young Lee, Hailin Jin*
2. [0835] SiamRPN++: Evolution of Siamese Visual Tracking With Very Deep Networks, *Bo Li, Wei Wu, Qiang Wang, Fangyi Zhang, Junliang Xing, Junjie Yan*
3. [0840] Sphere Generative Adversarial Network Based on Geometric Moment Matching, *Sung Woo Park, Junseok Kwon*
4. [0848] Adversarial Attacks Beyond the Image Space, *Xiaohui Zeng, Chenxi Liu, Yu-Siang Wang, Weichao Qiu, Lingxi Xie, Yu-Wing Tai, Chi-Keung Tang, Alan L. Yuille*
5. [0853] Evading Defenses to Transferable Adversarial Examples by Translation-Invariant Attacks, *Yinpeng Dong, Tianyu Pang, Hang Su, Jun Zhu*
6. [0858] Decoupling Direction and Norm for Efficient Gradient-Based L2 Adversarial Attacks and Defenses, *Jérôme Rony, Luiz G. Hafemann, Luiz S. Oliveira, Ismail Ben Ayed, Robert Sabourin, Eric Granger*
7. [0906] A General and Adaptive Robust Loss Function, *Jonathan T. Barron*
8. [0911] Filter Pruning via Geometric Median for Deep Convolutional Neural Networks Acceleration, *Yang He, Ping Liu, Ziwei Wang, Zhilan Hu, Yi Yang*
9. [0916] Learning to Quantize Deep Networks by Optimizing Quantization Intervals With Task Loss, *Sangil Jung, Changyong Son, Seohyung Lee, Jinwoo Son, Jae-Joon Han, Youngjun Kwak, Sung Ju Hwang, Changkyu Choi*
10. [0924] Not All Areas Are Equal: Transfer Learning for Semantic Segmentation via Hierarchical Region Selection, *Ruoqi Sun, Xinge Zhu, Chongruo Wu, Chen Huang, Jianping Shi, Lizhuang Ma*
11. [0929] Unsupervised Learning of Dense Shape Correspondence, *Oshri Halimi, Or Litany, Emanuele Rodolà, Alex M. Bronstein, Ron Kimmel*
12. [0934] Unsupervised Visual Domain Adaptation: A Deep Max-Margin Gaussian Process Approach, *Minyoung Kim, Pritish Sahu, Behnam Gholami, Vladimir Pavlovic*
13. [0942] Balanced Self-Paced Learning for Generative Adversarial Clustering Network, *Kamran Ghasedi, Xiaoqian Wang, Cheng Deng, Heng Huang*
14. [0947] A Style-Based Generator Architecture for Generative Adversarial Networks, *Tero Karras, Samuli Laine, Timo Aila*
15. [0952] Parallel Optimal Transport GAN, *Gil Avraham, Yan Zuo, Tom Drummond*

0830-1000 Oral Session 2-1B: 3D Single View & RGBD
(Grand Ballroom)

Papers in this session are in Poster Session 2-1P (Posters 106-120)

Chairs: David Fouhey (*Univ. of Michigan*)
Saurabh Gupta (*Facebook AI Research; UIUC*)

Format (5 min. presentation; 3 min. group questions/3 papers)

1. [0830] 3D-SIS: 3D Semantic Instance Segmentation of RGB-D Scans, *Ji Hou, Angela Dai, Matthias Nießner*
2. [0835] Causes and Corrections for Bimodal Multi-Path Scanning With Structured Light, *Yu Zhang, Daniel L. Lau, Ying Yu*
3. [0840] TextureNet: Consistent Local Parametrizations for Learning From High-Resolution Signals on Meshes, *Jingwei Huang, Haotian Zhang, Li Yi, Thomas Funkhouser, Matthias Nießner, Leonidas J. Guibas*
4. [0848] PlaneRCNN: 3D Plane Detection and Reconstruction From a Single Image, *Chen Liu, Kihwan Kim, Jinwei Gu, Yasutaka Furukawa, Jan Kautz*
5. [0853] Occupancy Networks: Learning 3D Reconstruction in Function Space, *Lars Mescheder, Michael Oechsle, Michael Niemeyer, Sebastian Nowozin, Andreas Geiger*
6. [0858] 3D Shape Reconstruction From Images in the Frequency Domain, *Weichao Shen, Yunde Jia, Yuwei Wu*
7. [0906] SiCloPe: Silhouette-Based Clothed People, *Ryota Natsume, Shunsuke Saito, Zeng Huang, Weikai Chen, Chongyang Ma, Hao Li, Shigeo Morishima*
8. [0911] Detailed Human Shape Estimation From a Single Image by Hierarchical Mesh Deformation, *Hao Zhu, Xinxin Zuo, Sen Wang, Xun Cao, Ruigang Yang*
9. [0916] Convolutional Mesh Regression for Single-Image Human Shape Reconstruction, *Nikos Kolotouros, Georgios Pavlakos, Kostas Daniilidis*
10. [0924] H+O: Unified Egocentric Recognition of 3D Hand-Object Poses and Interactions, *Bugra Tekin, Federica Bogo, Marc Pollefeys*
11. [0929] Learning the Depths of Moving People by Watching Frozen People, *Zhengqi Li, Tali Dekel, Forrester Cole, Richard Tucker, Noah Snavely, Ce Liu, William T. Freeman*
12. [0934] Extreme Relative Pose Estimation for RGB-D Scans via Scene Completion, *Zhenpei Yang, Jeffrey Z. Pan, Linjie Luo, Xiaowei Zhou, Kristen Grauman, Qixing Huang*
13. [0942] A Skeleton-Bridged Deep Learning Approach for Generating Meshes of Complex Topologies From Single RGB Images, *Jiapeng Tang, Xiaoguang Han, Junyi Pan, Kui Jia, Xin Tong*
14. [0947] Learning Structure-And-Motion-Aware Rolling Shutter Correction, *Bingbing Zhuang, Quoc-Huy Tran, Pan Ji, Loong-Fah Cheong, Manmohan Chandraker*
15. [0952] PVNet: Pixel-Wise Voting Network for 6DoF Pose Estimation, *Sida Peng, Yuan Liu, Qixing Huang, Xiaowei Zhou, Hujun Bao*

0830–1000 Oral Session 2-1C: Motion & Biometrics
(Promenade Ballroom)

Papers in this session are in Poster Session 2-1P (Posters 135–149)

Chairs: Jia-Bin Huang (*Virginia Tech*)
Ajay Kumar (*Hong Kong Polytechnic Univ.*)

Format (5 min. presentation; 3 min. group questions/3 papers)

1. [0830] SelfFlow: Self-Supervised Learning of Optical Flow, Pengpeng Liu, Michael Lyu, Irwin King, Jia Xu
2. [0835] Taking a Deeper Look at the Inverse Compositional Algorithm, Zhaoyang Lv, Frank Dellaert, James M. Rehg, Andreas Geiger
3. [0840] Deeper and Wider Siamese Networks for Real-Time Visual Tracking, Zhipeng Zhang, Houwen Peng
4. [0848] Self-Supervised Adaptation of High-Fidelity Face Models for Monocular Performance Tracking, Jae Shin Yoon, Takaaki Shiratori, Shoou-I Yu, Hyun Soo Park
5. [0853] Diverse Generation for Multi-Agent Sports Games, Raymond A. Yeh, Alexander G. Schwing, Jonathan Huang, Kevin Murphy
6. [0858] Efficient Online Multi-Person 2D Pose Tracking With Recurrent Spatio-Temporal Affinity Fields, Yaadhav Raaj, Haroon Idrees, Gines Hidalgo, Yaser Sheikh
7. [0906] GFrames: Gradient-Based Local Reference Frame for 3D Shape Matching, Simone Melzi, Riccardo Spezialetti, Federico Tombari, Michael M. Bronstein, Luigi Di Stefano, Emanuele Rodolà
8. [0911] Eliminating Exposure Bias and Metric Mismatch in Multiple Object Tracking, Andrii Maksai, Pascal Fua
9. [0916] Graph Convolutional Tracking, Junyu Gao, Tianzhu Zhang, Changsheng Xu
10. [0924] ATOM: Accurate Tracking by Overlap Maximization, Martin Danelljan, Goutam Bhat, Fahad Shahbaz Khan, Michael Felsberg
11. [0929] Visual Tracking via Adaptive Spatially-Regularized Correlation Filters, Kenan Dai, Dong Wang, Huchuan Lu, Chong Sun, Jianhua Li
12. [0934] Deep Tree Learning for Zero-Shot Face Anti-Spoofing, Yaojie Liu, Joel Stehouwer, Amin Jourabloo, Xiaoming Liu
13. [0942] ArcFace: Additive Angular Margin Loss for Deep Face Recognition, Jiankang Deng, Jia Guo, Niannan Xue, Stefanos Zafeiriou
14. [0947] Learning Joint Gait Representation via Quintuplet Loss Minimization, Kaihao Zhang, Wenhan Luo, Lin Ma, Wei Liu, Hongdong Li
15. [0952] Gait Recognition via Disentangled Representation Learning, Ziyuan Zhang, Luan Tran, Xi Yin, Yousef Atoum, Xiaoming Liu, Jian Wan, Nanxin Wang

1000–1100 Morning Break (Exhibit Hall)

1000–1245 Demos (Exhibit Hall)

- Real Time Self-Adaptive Deep Stereo, Alessio Tonioni, Fabio Tosi, Matteo Poggi, Stefano Mattoccia, Luigi Di Stefano (*Univ. of Bologna*)
- XNect: Real-Time Multi-Person 3D Human Pose Estimation With a Single RGB Camera, Dushyant Mehta, Oleksandr Sotnychenko, Franziska Mueller, Weipeng Xu, Hans-Peter Seidel, Pascal Fua, Mohamed Elgharib, Helge Rhodin, Gerard Pons-Moll, Christian Theobalt (*Max Planck Institute for Informatics*)

hamed Elgharib, Helge Rhodin, Gerard Pons-Moll, Christian Theobalt (*Max Planck Institute for Informatics*)

- CRAVES: Controlling Robotic Arm With a Vision-Based Economic System, Yiming Zuo, Weichao Qiu, Lingxi Xie, Fangwei Zhong, Yizhou Wang, Alan Yuille (*Johns Hopkins Univ.*)
- HoloPose: Holistic 3D Human Body Estimation In-the-Wild, Riza Alp Guler, George Papandreou, Stefanos Zafeiriou, Iasonas Kokkinos (*Ariel AI*)

1000–1245 Exhibits (Exhibit Hall)

- See Exhibits map for list of exhibitors.

1000–1245 Poster Session 2-1P (Exhibit Hall)

Deep Learning

1. Learning Video Representations From Correspondence Proposals, Xingyu Liu, Joon-Young Lee, Hailin Jin
2. SiamRPN++: Evolution of Siamese Visual Tracking With Very Deep Networks, Bo Li, Wei Wu, Qiang Wang, Fangyi Zhang, Junliang Xing, Junjie Yan
3. Sphere Generative Adversarial Network Based on Geometric Moment Matching, Sung Woo Park, Junseok Kwon
4. Adversarial Attacks Beyond the Image Space, Xiaohui Zeng, Chenxi Liu, Yu-Siang Wang, Weichao Qiu, Lingxi Xie, Yu-Wing Tai, Chi-Keung Tang, Alan L. Yuille
5. Evading Defenses to Transferable Adversarial Examples by Translation-Invariant Attacks, Yinpeng Dong, Tianyu Pang, Hang Su, Jun Zhu
6. Decoupling Direction and Norm for Efficient Gradient-Based L2 Adversarial Attacks and Defenses, Jérôme Rony, Luiz G. Hafemann, Luiz S. Oliveira, Ismail Ben Ayed, Robert Sabourin, Eric Granger
7. A General and Adaptive Robust Loss Function, Jonathan T. Barron
8. Filter Pruning via Geometric Median for Deep Convolutional Neural Networks Acceleration, Yang He, Ping Liu, Ziwei Wang, Zhilan Hu, Yi Yang
9. Learning to Quantize Deep Networks by Optimizing Quantization Intervals With Task Loss, Sangil Jung, Changyong Son, Seohyung Lee, Jinwoo Son, Jae-Joon Han, Youngjun Kwak, Sung Ju Hwang, Changkyu Choi
10. Not All Areas Are Equal: Transfer Learning for Semantic Segmentation via Hierarchical Region Selection, Ruoqi Sun, Xinge Zhu, Congruo Wu, Chen Huang, Jianping Shi, Lizhuang Ma
11. Unsupervised Learning of Dense Shape Correspondence, Oshri Halimi, Or Litany, Emanuele Rodolà, Alex M. Bronstein, Ron Kimmel
12. Unsupervised Visual Domain Adaptation: A Deep Max-Margin Gaussian Process Approach, Minyoung Kim, Pritish Sahu, Behnam Gholami, Vladimir Pavlovic
13. Balanced Self-Paced Learning for Generative Adversarial Clustering Network, Kamran Ghasedi, Xiaoqian Wang, Cheng Deng, Heng Huang
14. A Style-Based Generator Architecture for Generative Adversarial Networks, Tero Karras, Samuli Laine, Timo Aila
15. Parallel Optimal Transport GAN, Gil Avraham, Yan Zuo, Tom Drummond
16. Reversible GANs for Memory-Efficient Image-To-Image Translation, Tycho F.A. van der Ouderaa, Daniel E. Worrall

17. Sensitive-Sample Fingerprinting of Deep Neural Networks, *Zecheng He, Tianwei Zhang, Ruby Lee*
 18. Soft Labels for Ordinal Regression, *Raúl Díaz, Amit Marathe*
 19. Local to Global Learning: Gradually Adding Classes for Training Deep Neural Networks, *Hao Cheng, Dongze Lian, Bowen Deng, Shenghua Gao, Tao Tan, Yanlin Geng*
 20. What Does It Mean to Learn in Deep Networks? And, How Does One Detect Adversarial Attacks? *Ciprian A. Corneanu, Meysam Madadi, Sergio Escalera, Aleix M. Martinez*
 21. Handwriting Recognition in Low-Resource Scripts Using Adversarial Learning, *Ayan Kumar Bhunia, Abhirup Das, Ankan Kumar Bhunia, Perla Sai Raj Kishore, Partha Pratim Roy*
 22. Adversarial Defense Through Network Profiling Based Path Extraction, *Yuxian Qiu, Jingwen Leng, Cong Guo, Quan Chen, Chao Li, Minyi Guo, Yuhao Zhu*
 23. RENAS: Reinforced Evolutionary Neural Architecture Search, *Yukang Chen, Gaofeng Meng, Qian Zhang, Shiming Xiang, Chang Huang, Lisen Mu, Xinggang Wang*
 24. Co-Occurrence Neural Network, *Irina Shevlev, Shai Avidan*
 25. SpotTune: Transfer Learning Through Adaptive Fine-Tuning, *Yunhui Guo, Honghui Shi, Abhishek Kumar, Kristen Grauman, Tajana Rosing, Rogerio Feris*
 26. Signal-To-Noise Ratio: A Robust Distance Metric for Deep Metric Learning, *Tongtong Yuan, Weihong Deng, Jian Tang, Yinan Tang, Binghui Chen*
 27. Detection Based Defense Against Adversarial Examples From the Steganalysis Point of View, *Jiayang Liu, Weiming Zhang, Yiwei Zhang, Dongdong Hou, Yujia Liu, Hongyue Zha, Nenghai Yu*
 28. HetConv: Heterogeneous Kernel-Based Convolutions for Deep CNNs, *Pravendra Singh, Vinay Kumar Verma, Piyush Rai, Vinay P. Namboodiri*
 29. Strike (With) a Pose: Neural Networks Are Easily Fooled by Strange Poses of Familiar Objects, *Michael A. Alcorn, Qi Li, Zhitao Gong, Chengfei Wang, Long Mai, Wei-Shinn Ku, Anh Nguyen*
 30. Blind Geometric Distortion Correction on Images Through Deep Learning, *Xiaoyu Li, Bo Zhang, Pedro V. Sander, Jing Liao*
 31. Instance-Level Meta Normalization, *Songhao Jia, Ding-Jie Chen, Hwann-Tzong Chen*
 32. Iterative Normalization: Beyond Standardization Towards Efficient Whitening, *Lei Huang, Yi Zhou, Fan Zhu, Li Liu, Ling Shao*
 33. On Learning Density Aware Embeddings, *Soumyadeep Ghosh, Richa Singh, Mayank Vatsa*
 34. Contrastive Adaptation Network for Unsupervised Domain Adaptation, *Guoliang Kang, Lu Jiang, Yi Yang, Alexander G. Hauptmann*
 35. LP-3DCNN: Unveiling Local Phase in 3D Convolutional Neural Networks, *Sudhakar Kumawat, Shanmuganathan Raman*
 36. Attribute-Driven Feature Disentangling and Temporal Aggregation for Video Person Re-Identification, *Yiru Zhao, Xu Shen, Zhongming Jin, Hongtao Lu, Xian-sheng Hua*
 37. Binary Ensemble Neural Network: More Bits per Network or More Networks per Bit? *Shilin Zhu, Xin Dong, Hao Su*
 38. Distilling Object Detectors With Fine-Grained Feature Imitation, *Tao Wang, Li Yuan, Xiaopeng Zhang, Jiashi Feng*
 39. Centripetal SGD for Pruning Very Deep Convolutional Networks With Complicated Structure, *Xiaohan Ding, Guiguang Ding, Yuchen Guo, Jungong Han*
 40. Knockoff Nets: Stealing Functionality of Black-Box Models, *Tribhuvanesh Orekondy, Bernt Schiele, Mario Fritz*
- Recognition**
41. Deep Embedding Learning With Discriminative Sampling Policy, *Yueqi Duan, Lei Chen, Jiwen Lu, Jie Zhou*
 42. Hybrid Task Cascade for Instance Segmentation, *Kai Chen, Jiangmiao Pang, Jiaqi Wang, Yu Xiong, Xiaoxiao Li, Shuyang Sun, Wansen Feng, Ziwei Liu, Jianping Shi, Wanli Ouyang, Chen Change Loy, Dahua Lin*
 43. Multi-Task Self-Supervised Object Detection via Recycling of Bounding Box Annotations, *Wonhee Lee, Joonil Na, Gunhee Kim*
 44. ClusterNet: Deep Hierarchical Cluster Network With Rigorously Rotation-Invariant Representation for Point Cloud Analysis, *Chao Chen, Guanbin Li, Ruijia Xu, Tianshui Chen, Meng Wang, Liang Lin*
 45. Learning to Learn Relation for Important People Detection in Still Images, *Wei-Hong Li, Fa-Ting Hong, Wei-Shi Zheng*
 46. Looking for the Devil in the Details: Learning Trilinear Attention Sampling Network for Fine-Grained Image Recognition, *Heliang Zheng, Jianlong Fu, Zheng-Jun Zha, Jiebo Luo*
 47. Multi-Similarity Loss With General Pair Weighting for Deep Metric Learning, *Xun Wang, Xintong Han, Weilin Huang, Dengke Dong, Matthew R. Scott*
 48. Domain-Symmetric Networks for Adversarial Domain Adaptation, *Yabin Zhang, Hui Tang, Kui Jia, Mingkui Tan*
 49. End-To-End Supervised Product Quantization for Image Search and Retrieval, *Benjamin Klein, Lior Wolf*
 50. Learning to Learn From Noisy Labeled Data, *Junnan Li, Yongkang Wong, Qi Zhao, Mohan S. Kankanhalli*
 51. DSFD: Dual Shot Face Detector, *Jian Li, Yabiao Wang, Changan Wang, Ying Tai, Jianjun Qian, Jian Yang, Chengjie Wang, Jilin Li, Feiyue Huang*
 52. Label Propagation for Deep Semi-Supervised Learning, *Ahmet Iscen, Giorgos Tolias, Yannis Avrithis, Ondřej Chum*
 53. Deep Global Generalized Gaussian Networks, *Qilong Wang, Peihua Li, Qinghua Hu, Pengfei Zhu, Wangmeng Zuo*
 54. Semantically Tied Paired Cycle Consistency for Zero-Shot Sketch-Based Image Retrieval, *Anjan Dutta, Zeynep Akata*
 55. Context-Aware Crowd Counting, *Weizhe Liu, Mathieu Salzmann, Pascal Fua*
 56. Detect-To-Retrieve: Efficient Regional Aggregation for Image Search, *Marvin Teichmann, André Araujo, Menglong Zhu, Jack Sim*
 57. Towards Accurate One-Stage Object Detection With AP-Loss, *Kean Chen, Jianguo Li, Weiyao Lin, John See, Ji Wang, Lingyu Duan, Zhibo Chen, Changwei He, Junni Zou*
 58. On Exploring Undetermined Relationships for Visual Relationship Detection, *Yibing Zhan, Jun Yu, Ting Yu, Dacheng Tao*
 59. Learning Without Memorizing, *Prithviraj Dhar, Rajat Vikram Singh, Kuan-Chuan Peng, Ziyang Wu, Rama Chellappa*
 60. Dynamic Recursive Neural Network, *Qiushan Guo, Zhipeng Yu, Yichao Wu, Ding Liang, Haoyu Qin, Junjie Yan*
 61. Destruction and Construction Learning for Fine-Grained Image Recognition, *Yue Chen, Yalong Bai, Wei Zhang, Tao Mei*
 62. Distraction-Aware Shadow Detection, *Quanlong Zheng, Xiaotian Qiao, Ying Cao, Rynson W.H. Lau*
 63. Multi-Label Image Recognition With Graph Convolutional Networks, *Zhao-Min Chen, Xiu-Shen Wei, Peng Wang, Yanwen Guo*

64. High-Level Semantic Feature Detection: A New Perspective for Pedestrian Detection, *Wei Liu, Shengcai Liao, Weiqiang Ren, Weidong Hu, Yinan Yu*
65. RepMet: Representative-Based Metric Learning for Classification and Few-Shot Object Detection, *Leonid Karlinsky, Joseph Shtok, Sivan Harary, Eli Schwartz, Amit Aides, Rogerio Feris, Raja Giryes, Alex M. Bronstein*
66. Ranked List Loss for Deep Metric Learning, *Xinshao Wang, Yang Hua, Elyor Kodirov, Guosheng Hu, Romain Garnier, Neil M. Robertson*
67. CANet: Class-Agnostic Segmentation Networks With Iterative Refinement and Attentive Few-Shot Learning, *Chi Zhang, Guosheng Lin, Fayao Liu, Rui Yao, Chunhua Shen*
68. Precise Detection in Densely Packed Scenes, *Eran Goldman, Roei Herzig, Aviv Eisenschat, Jacob Goldberger, Tal Hassner*

Segmentation, Grouping, & Shape

69. KE-GAN: Knowledge Embedded Generative Adversarial Networks for Semi-Supervised Scene Parsing, *Mengshi Qi, Yunhong Wang, Jie Qin, Annan Li*
70. Fast User-Guided Video Object Segmentation by Interaction-And-Propagation Networks, *Seoung Wug Oh, Joon-Young Lee, Ning Xu, Seon Joo Kim*
71. Fast Interactive Object Annotation With Curve-GCN, *Huan Ling, Jun Gao, Amlan Kar, Wenzheng Chen, Sanja Fidler*
72. FickleNet: Weakly and Semi-Supervised Semantic Image Segmentation Using Stochastic Inference, *Jungbeom Lee, Eunji Kim, Sungmin Lee, Jangho Lee, Sungho Yoon*
73. RVOS: End-To-End Recurrent Network for Video Object Segmentation, *Carles Ventura, Miriam Bellver, Andreu Girbau, Amaia Salvador, Ferran Marques, Xavier Giro-i-Nieto*
74. DeepFlux for Skeletons in the Wild, *Yukang Wang, Yongchao Xu, Stavros Tsogkas, Xiang Bai, Sven Dickinson, Kaleem Siddiqi*
75. Interactive Image Segmentation via Backpropagating Refinement Scheme, *Won-Dong Jang, Chang-Su Kim*
76. Scene Parsing via Integrated Classification Model and Variance-Based Regularization, *Hengcan Shi, Hongliang Li, Qingbo Wu, Zichen Song*

Statistics, Physics, Theory, & Datasets

77. RAVEN: A Dataset for Relational and Analogical Visual REasoning, *Chi Zhang, Feng Gao, Baoxiong Jia, Yixin Zhu, Song-Chun Zhu*
78. Surface Reconstruction From Normals: A Robust DGP-Based Discontinuity Preservation Approach, *Wuyuan Xie, Miaohui Wang, Mingqiang Wei, Jianmin Jiang, Jing Qin*
79. DeepFashion2: A Versatile Benchmark for Detection, Pose Estimation, Segmentation and Re-Identification of Clothing Images, *Yuying Ge, Ruimao Zhang, Xiaogang Wang, Xiaoou Tang, Ping Luo*
80. Jumping Manifolds: Geometry Aware Dense Non-Rigid Structure From Motion, *Suryansh Kumar*
81. LVIS: A Dataset for Large Vocabulary Instance Segmentation, *Agrim Gupta, Piotr Dollár, Ross Girshick*
82. Fast Object Class Labelling via Speech, *Michael Gygli, Vittorio Ferrari*
83. LaSOT: A High-Quality Benchmark for Large-Scale Single Object Tracking, *Heng Fan, Liting Lin, Fan Yang, Peng Chu, Ge Deng, Sijia Yu, Hexin Bai, Yong Xu, Chunyuan Liao, Haibin Ling*
84. Creative Flow+ Dataset, *Maria Shugrina, Ziheng Liang, Amlan Kar, Jiaman Li, Angad Singh, Karan Singh, Sanja Fidler*
85. Weakly Supervised Open-Set Domain Adaptation by Dual-Domain Collaboration, *Shuhan Tan, Jiening Jiao, Wei-Shi Zheng*
86. A Neurobiological Evaluation Metric for Neural Network Model Search, *Nathaniel Blanchard, Jeffery Kinnison, Brandon RichardWebster, Pouya Bashivan, Walter J. Scheirer*
87. Iterative Projection and Matching: Finding Structure-Preserving Representatives and Its Application to Computer Vision, *Alireza Zaeemzadeh, Mohsen Joneidi, Nazanin Rahnavard, Mubarak Shah*
88. Efficient Multi-Domain Learning by Covariance Normalization, *Yunsheng Li, Nuno Vasconcelos*
89. Predicting Visible Image Differences Under Varying Display Brightness and Viewing Distance, *Nanyang Ye, Krzysztof Wolski, Rafał K. Mantiuk*
90. A Bayesian Perspective on the Deep Image Prior, *Zezhou Cheng, Matheus Gadelha, Subhransu Maji, Daniel Sheldon*
91. ApolloCar3D: A Large 3D Car Instance Understanding Benchmark for Autonomous Driving, *Xibin Song, Peng Wang, Dingfu Zhou, Rui Zhu, Chenye Guan, Yuchao Dai, Hao Su, Hongdong Li, Ruigang Yang*
92. Compressing Unknown Images With Product Quantizer for Efficient Zero-Shot Classification, *Jin Li, Xuguang Lan, Yang Liu, Le Wang, Nanning Zheng*
93. Self-Supervised Convolutional Subspace Clustering Network, *Junjian Zhang, Chun-Guang Li, Chong You, Xianbiao Qi, Honggang Zhang, Jun Guo, Zhouchen Lin*

3D Multiview

94. Multi-Scale Geometric Consistency Guided Multi-View Stereo, *Qingshan Xu, Wenbing Tao*
95. Privacy Preserving Image-Based Localization, *Pablo Speciale, Johannes L. Schönberger, Sing Bing Kang, Sudipta N. Sinha, Marc Pollefeys*
96. SimulCap : Single-View Human Performance Capture With Cloth Simulation, *Tao Yu, Zerong Zheng, Yuan Zhong, Jianhui Zhao, Qionghai Dai, Gerard Pons-Moll, Yebin Liu*
97. Hierarchical Deep Stereo Matching on High-Resolution Images, *Gengshan Yang, Joshua Manela, Michael Happold, Deva Ramanan*
98. Recurrent MVSNet for High-Resolution Multi-View Stereo Depth Inference, *Yao Yao, Zixin Luo, Shiwei Li, Tianwei Shen, Tian Fang, Long Quan*
99. Synthesizing 3D Shapes From Silhouette Image Collections Using Multi-Projection Generative Adversarial Networks, *Xiao Li, Yue Dong, Pieter Peers, Xin Tong*
100. The Perfect Match: 3D Point Cloud Matching With Smoothed Densities, *Zan Gojcic, Caifa Zhou, Jan D. Wegner, Andreas Wieser*
101. Recurrent Neural Network for (Un-)Supervised Learning of Monocular Video Visual Odometry and Depth, *Rui Wang, Stephen M. Pizer, Jan-Michael Frahm*
102. PointWeb: Enhancing Local Neighborhood Features for Point Cloud Processing, *Hengshuang Zhao, Li Jiang, Chi-Wing Fu, Jiaya Jia*
103. Scan2Mesh: From Unstructured Range Scans to 3D Meshes, *Angela Dai, Matthias Nießner*
104. Unsupervised Domain Adaptation for ToF Data Denoising With Adversarial Learning, *Gianluca Agresti, Henrik Schaefer, Piergiorgio Sartor, Pietro Zanuttigh*
105. Learning Independent Object Motion From Unlabelled Stereoscopic Videos, *Zhe Cao, Abhishek Kar, Christian Häne, Jitendra Malik*

3D Single View & RGBD

106. 3D-SIS: 3D Semantic Instance Segmentation of RGB-D Scans, *Ji Hou, Angela Dai, Matthias Nießner*
107. Causes and Corrections for Bimodal Multi-Path Scanning With Structured Light, *Yu Zhang, Daniel L. Lau, Ying Yu*
108. TextureNet: Consistent Local Parametrizations for Learning From High-Resolution Signals on Meshes, *Jingwei Huang, Haotian Zhang, Li Yi, Thomas Funkhouser, Matthias Nießner, Leonidas J. Guibas*
109. PlaneRCNN: 3D Plane Detection and Reconstruction From a Single Image, *Chen Liu, Kihwan Kim, Jinwei Gu, Yasutaka Furukawa, Jan Kautz*
110. Occupancy Networks: Learning 3D Reconstruction in Function Space, *Lars Mescheder, Michael Oechsle, Michael Niemeyer, Sebastian Nowozin, Andreas Geiger*
111. 3D Shape Reconstruction From Images in the Frequency Domain, *Weichao Shen, Yunde Jia, Yuwei Wu*
112. SiCloPe: Silhouette-Based Clothed People, *Ryota Natsume, Shunsuke Saito, Zeng Huang, Weikai Chen, Chongyang Ma, Hao Li, Shigeo Morishima*
113. Detailed Human Shape Estimation From a Single Image by Hierarchical Mesh Deformation, *Hao Zhu, Xinxin Zuo, Sen Wang, Xun Cao, Ruigang Yang*
114. Convolutional Mesh Regression for Single-Image Human Shape Reconstruction, *Nikos Kolotouros, Georgios Pavlakos, Kostas Daniilidis*
115. H+O: Unified Egocentric Recognition of 3D Hand-Object Poses and Interactions, *Bugra Tekin, Federica Bogo, Marc Pollefeys*
116. Learning the Depths of Moving People by Watching Frozen People, *Zhengqi Li, Tali Dekel, Forrester Cole, Richard Tucker, Noah Snavely, Ce Liu, William T. Freeman*
117. Extreme Relative Pose Estimation for RGB-D Scans via Scene Completion, *Zhenpei Yang, Jeffrey Z. Pan, Linjie Luo, Xiaowei Zhou, Kristen Grauman, Qixing Huang*
118. A Skeleton-Bridged Deep Learning Approach for Generating Meshes of Complex Topologies From Single RGB Images, *Jiapeng Tang, Xiaoguang Han, Junyi Pan, Kui Jia, Xin Tong*
119. Learning Structure-And-Motion-Aware Rolling Shutter Correction, *Bingbing Zhuang, Quoc-Huy Tran, Pan Ji, Loong-Fah Cheong, Manmohan Chandraker*
120. PVNet: Pixel-Wise Voting Network for 6DoF Pose Estimation, *Sida Peng, Yuan Liu, Qixing Huang, Xiaowei Zhou, Hujun Bao*
121. Learning Single-Image Depth From Videos Using Quality Assessment Networks, *Weifeng Chen, Shengyi Qian, Jia Deng*
122. Learning 3D Human Dynamics From Video, *Angjoo Kanazawa, Jason Y. Zhang, Panna Felsen, Jitendra Malik*
123. Lending Orientation to Neural Networks for Cross-View Geo-Localization, *Liu Liu, Hongdong Li*
124. Visual Localization by Learning Objects-Of-Interest Dense Match Regression, *Philippe Weinzaepfel, Gabriela Csurka, Yohann Cabon, Martin Humenberger*
125. Bilateral Cyclic Constraint and Adaptive Regularization for Unsupervised Monocular Depth Prediction, *Alex Wong, Stefano Soatto*

Face & Body

126. Face Parsing With Rol Tanh-Warping, *Jinpeng Lin, Hao Yang, Dong Chen, Ming Zeng, Fang Wen, Lu Yuan*
127. Multi-Person Articulated Tracking With Spatial and Temporal Embeddings, *Sheng Jin, Wentao Liu, Wanli Ouyang, Chen Qian*

128. Multi-Person Pose Estimation With Enhanced Channel-Wise and Spatial Information, *Kai Su, Dongdong Yu, Zhenqi Xu, Xin Geng, Changhu Wang*
129. A Compact Embedding for Facial Expression Similarity, *Raviteja Vemulapalli, Aseem Agarwala*
130. Deep High-Resolution Representation Learning for Human Pose Estimation, *Ke Sun, Bin Xiao, Dong Liu, Jingdong Wang*
131. Feature Transfer Learning for Face Recognition With Under-Represented Data, *Xi Yin, Xiang Yu, Kihyuk Sohn, Xiaoming Liu, Manmohan Chandraker*
132. Unsupervised 3D Pose Estimation With Geometric Self-Supervision, *Ching-Hang Chen, Amrith Tyagi, Amit Agrawal, Dylan Drover, Rohith MV, Stefan Stojanov, James M. Rehg*

Action & Video

133. Peeking Into the Future: Predicting Future Person Activities and Locations in Videos, *Junwei Liang, Lu Jiang, Juan Carlos Niebles, Alexander G. Hauptmann, Li Fei-Fei*
134. Re-Identification With Consistent Attentive Siamese Networks, *Meng Zheng, Srikrishna Karanam, Ziyang Wu, Richard J. Radke*

Motion & Biometrics

135. SelfFlow: Self-Supervised Learning of Optical Flow, *Pengpeng Liu, Michael Lyu, Irwin King, Jia Xu*
136. Taking a Deeper Look at the Inverse Compositional Algorithm, *Zhaoyang Lv, Frank Dellaert, James M. Rehg, Andreas Geiger*
137. Deeper and Wider Siamese Networks for Real-Time Visual Tracking, *Zhipeng Zhang, Houwen Peng*
138. Self-Supervised Adaptation of High-Fidelity Face Models for Monocular Performance Tracking, *Jae Shin Yoon, Takaaki Shiratori, Shoou-I Yu, Hyun Soo Park*
139. Diverse Generation for Multi-Agent Sports Games, *Raymond A. Yeh, Alexander G. Schwing, Jonathan Huang, Kevin Murphy*
140. Efficient Online Multi-Person 2D Pose Tracking With Recurrent Spatio-Temporal Affinity Fields, *Yaadhav Raaj, Haroon Idrees, Gines Hidalgo, Yaser Sheikh*
141. GFrames: Gradient-Based Local Reference Frame for 3D Shape Matching, *Simone Melzi, Riccardo Spezialetti, Federico Tombari, Michael M. Bronstein, Luigi Di Stefano, Emanuele Rodolà*
142. Eliminating Exposure Bias and Metric Mismatch in Multiple Object Tracking, *Andrii Maksai, Pascal Fua*
143. Graph Convolutional Tracking, *Junyu Gao, Tianzhu Zhang, Changsheng Xu*
144. ATOM: Accurate Tracking by Overlap Maximization, *Martin Danelljan, Goutam Bhat, Fahad Shahbaz Khan, Michael Felsberg*
145. Visual Tracking via Adaptive Spatially-Regularized Correlation Filters, *Kenan Dai, Dong Wang, Huchuan Lu, Chong Sun, Jianhua Li*
146. Deep Tree Learning for Zero-Shot Face Anti-Spoofing, *Yaojie Liu, Joel Stehouwer, Amin Jourabloo, Xiaoming Liu*
147. ArcFace: Additive Angular Margin Loss for Deep Face Recognition, *Jiankang Deng, Jia Guo, Niannan Xue, Stefanos Zafeiriou*
148. Learning Joint Gait Representation via Quintuplet Loss Minimization, *Kaihao Zhang, Wenhan Luo, Lin Ma, Wei Liu, Hongdong Li*
149. Gait Recognition via Disentangled Representation Learning, *Ziyuan Zhang, Luan Tran, Xi Yin, Yusef Atoum, Xiaoming Liu, Jian Wan, Nanxin Wang*

150. On the Continuity of Rotation Representations in Neural Networks, *Yi Zhou, Connelly Barnes, Jingwan Lu, Jimei Yang, Hao Li*
151. Iterative Residual Refinement for Joint Optical Flow and Occlusion Estimation, *Junhwa Hur, Stefan Roth*
152. Inverse Discriminative Networks for Handwritten Signature Verification, *Ping Wei, Huan Li, Ping Hu*
153. Led3D: A Lightweight and Efficient Deep Approach to Recognizing Low-Quality 3D Faces, *Guodong Mu, Di Huang, Guosheng Hu, Jia Sun, Yunhong Wang*
154. ROI Pooled Correlation Filters for Visual Tracking, *Yuxuan Sun, Chong Sun, Dong Wang, You He, Huchuan Lu*

Synthesis

155. Deep Video Inpainting, *Dahun Kim, Sanghyun Woo, Joon-Young Lee, In So Kweon*
156. DM-GAN: Dynamic Memory Generative Adversarial Networks for Text-To-Image Synthesis, *Minfeng Zhu, Pingbo Pan, Wei Chen, Yi Yang*
157. Non-Adversarial Image Synthesis With Generative Latent Nearest Neighbors, *Yedid Hoshen, Ke Li, Jitendra Malik*
158. Mixture Density Generative Adversarial Networks, *Hamid Eghbal-zadeh, Werner Zellinger, Gerhard Widmer*
159. SketchGAN: Joint Sketch Completion and Recognition With Generative Adversarial Network, *Fang Liu, Xiaoming Deng, Yu-Kun Lai, Yong-Jin Liu, Cuixia Ma, Hongan Wang*
160. Foreground-Aware Image Inpainting, *Wei Xiong, Jiahui Yu, Zhe Lin, Jimei Yang, Xin Lu, Connelly Barnes, Jiebo Luo*
161. Art2Real: Unfolding the Reality of Artworks via Semantically-Aware Image-To-Image Translation, *Matteo Tomei, Marcella Cornia, Lorenzo Baraldi, Rita Cucchiara*
162. Structure-Preserving Stereoscopic View Synthesis With Multi-Scale Adversarial Correlation Matching, *Yu Zhang, Dongqing Zou, Jimmy S. Ren, Zhe Jiang, Xiaohao Chen*
163. DynTypo: Example-Based Dynamic Text Effects Transfer, *Yifang Men, Zhouhui Lian, Yingmin Tang, Jianguo Xiao*
164. Arbitrary Style Transfer With Style-Attentional Networks, *Daeyeon Park, Kwang Hee Lee*
165. Typography With Decor: Intelligent Text Style Transfer, *Wenjing Wang, Jiaying Liu, Shuai Yang, Zongming Guo*
173. An Iterative and Cooperative Top-Down and Bottom-Up Inference Network for Salient Object Detection, *Wenguan Wang, Jianbing Shen, Ming-Ming Cheng, Ling Shao*
174. Deep Stacked Hierarchical Multi-Patch Network for Image Deblurring, *Hongguang Zhang, Yuchao Dai, Hongdong Li, Piotr Koniusz*
175. Turn a Silicon Camera Into an InGaAs Camera, *Feifan Lv, Yinqiang Zheng, Bohan Zhang, Feng Lu*
176. Low-Rank Tensor Completion With a New Tensor Nuclear Norm Induced by Invertible Linear Transforms, *Canyi Lu, Xi Peng, Yunchao Wei*
177. Joint Representative Selection and Feature Learning: A Semi-Supervised Approach, *Suchen Wang, Jingjing Meng, Junsong Yuan, Yap-Peng Tan*
178. The Domain Transform Solver, *Akash Bapat, Jan-Michael Frahm*
179. CapSal: Leveraging Captioning to Boost Semantics for Salient Object Detection, *Lu Zhang, Jianming Zhang, Zhe Lin, Huchuan Lu, You He*
180. Phase-Only Image Based Kernel Estimation for Single Image Blind Deblurring, *Liyuan Pan, Richard Hartley, Miaomiao Liu, Yuchao Dai*
181. Hierarchical Discrete Distribution Decomposition for Match Density Estimation, *Zhichao Yin, Trevor Darrell, Fisher Yu*
182. FOCNet: A Fractional Optimal Control Network for Image Denoising, *Xixi Jia, Sanyang Liu, Xiangchu Feng, Lei Zhang*
183. Orthogonal Decomposition Network for Pixel-Wise Binary Classification, *Chang Liu, Fang Wan, Wei Ke, Zhuowei Xiao, Yuan Yao, Xiaosong Zhang, Qixiang Ye*
184. Multi-Source Weak Supervision for Saliency Detection, *Yu Zeng, Yunzhi Zhuge, Huchuan Lu, Lihe Zhang, Mingyang Qian, Yizhou Yu*
185. ComDefend: An Efficient Image Compression Model to Defend Adversarial Examples, *XiaoJun Jia, Xingxing Wei, Xiaochun Cao, Hassan Foroosh*
186. Combinatorial Persistency Criteria for Multicut and Max-Cut, *Jan-Hendrik Lange, Bjoern Andres, Paul Swoboda*
187. S4Net: Single Stage Salient-Instance Segmentation, *Ruochen Fan, Ming-Ming Cheng, Qibin Hou, Tai-Jiang Mu, Jingdong Wang, Shi-Min Hu*
188. A Decomposition Algorithm for the Sparse Generalized Eigenvalue Problem, *Ganzhao Yuan, Li Shen, Wei-Shi Zheng*

Computational Photography & Graphics

166. RL-GAN-Net: A Reinforcement Learning Agent Controlled GAN Network for Real-Time Point Cloud Shape Completion, *Muhammad Sarmad, HyunJoo Jenny Lee, Young Min Kim*
167. Photo Wake-Up: 3D Character Animation From a Single Photo, *Chung-Yi Weng, Brian Curless, Ira Kemelmacher-Shlizerman*
168. DeepLight: Learning Illumination for Unconstrained Mobile Mixed Reality, *Chloe LeGendre, Wan-Chun Ma, Graham Fyffe, John Flynn, Laurent Charbonnel, Jay Busch, Paul Debevec*
169. Iterative Residual CNNs for Burst Photography Applications, *Filippos Kokkinos, Stamatis Lefkimmiatis*
170. Learning Implicit Fields for Generative Shape Modeling, *Zhiqin Chen, Hao Zhang*
171. Reliable and Efficient Image Cropping: A Grid Anchor Based Approach, *Hui Zeng, Lida Li, Zisheng Cao, Lei Zhang*
172. Patch-Based Progressive 3D Point Set Upsampling, *Wang Yifan, Shihao Wu, Hui Huang, Daniel Cohen-Or, Olga Sorkine-Hornung*

Low-Level & Optimization**Scenes & Representation**

189. Polynomial Representation for Persistence Diagram, *Zhichao Wang, Qian Li, Gang Li, Guandong Xu*
190. Crowd Counting and Density Estimation by Trellis Encoder-Decoder Networks, *Xiaolong Jiang, Zehao Xiao, Baochang Zhang, Xiantong Zhen, Xianbin Cao, David Doermann, Ling Shao*
191. Cross-Atlas Convolution for Parameterization Invariant Learning on Textured Mesh Surface, *Shiwei Li, Zixin Luo, Mingmin Zhen, Yao Yao, Tianwei Shen, Tian Fang, Long Quan*
192. Deep Surface Normal Estimation With Hierarchical RGB-D Fusion, *Jin Zeng, Yanfeng Tong, Yunmu Huang, Qiong Yan, Wenxiu Sun, Jing Chen, Yongtian Wang*
193. Knowledge-Embedded Routing Network for Scene Graph Generation, *Tianshui Chen, Weihao Yu, Riquan Chen, Liang Lin*
194. An End-To-End Network for Panoptic Segmentation, *HuanYu Liu, Chao Peng, Changqian Yu, Jingbo Wang, Xu Liu, Gang Yu, Wei Jiang*

195. Fast and Flexible Indoor Scene Synthesis via Deep Convolutional Generative Models, *Daniel Ritchie, Kai Wang, Yu-An Lin*
196. Marginalized Latent Semantic Encoder for Zero-Shot Learning, *Zhengming Ding, Hongfu Liu*
197. Scale-Adaptive Neural Dense Features: Learning via Hierarchical Context Aggregation, *Jaime Spencer, Richard Bowden, Simon Hadfield*
198. Unsupervised Embedding Learning via Invariant and Spreading Instance Feature, *Mang Ye, Xu Zhang, Pong C. Yuen, Shih-Fu Chang*
199. AOGNets: Compositional Grammatical Architectures for Deep Learning, *Xilai Li, Xi Song, Tianfu Wu*
200. A Robust Local Spectral Descriptor for Matching Non-Rigid Shapes With Incompatible Shape Structures, *Yiqun Wang, Jianwei Guo, Dong-Ming Yan, Kai Wang, Xiaopeng Zhang*

Language & Reasoning

201. Context and Attribute Grounded Dense Captioning, *Guojun Yin, Lu Sheng, Bin Liu, Nenghai Yu, Xiaogang Wang, Jing Shao*
202. Spot and Learn: A Maximum-Entropy Patch Sampler for Few-Shot Image Classification, *Wen-Hsuan Chu, Yu-Jhe Li, Jing-Cheng Chang, Yu-Chiang Frank Wang*
203. Interpreting CNNs via Decision Trees, *Quanshi Zhang, Yu Yang, Haotian Ma, Ying Nian Wu*
204. Dense Relational Captioning: Triple-Stream Networks for Relationship-Based Captioning, *Dong-Jin Kim, Jinsoo Choi, Tae-Hyun Oh, In So Kweon*
205. Deep Modular Co-Attention Networks for Visual Question Answering, *Zhou Yu, Jun Yu, Yuhao Cui, Dacheng Tao, Qi Tian*
206. Synthesizing Environment-Aware Activities via Activity Sketches, *Yuan-Hong Liao, Xavier Puig, Marko Boben, Antonio Torralba, Sanja Fidler*
207. Self-Critical n-Step Training for Image Captioning, *Junlong Gao, Shiqi Wang, Shanshe Wang, Siwei Ma, Wen Gao*
208. Multi-Target Embodied Question Answering, *Licheng Yu, Xinlei Chen, Georgia Gkioxari, Mohit Bansal, Tamara L. Berg, Dhruv Batra*
209. Visual Question Answering as Reading Comprehension, *Hui Li, Peng Wang, Chunhua Shen, Anton van den Hengel*
210. StoryGAN: A Sequential Conditional GAN for Story Visualization, *Yitong Li, Zhe Gan, Yelong Shen, Jingjing Liu, Yu Cheng, Yuexin Wu, Lawrence Carin, David Carlson, Jianfeng Gao*

Applications, Medical, & Robotics

211. Noise-Aware Unsupervised Deep Lidar-Stereo Fusion, *Xuelian Cheng, Yiran Zhong, Yuchao Dai, Pan Ji, Hongdong Li*
212. Versatile Multiple Choice Learning and Its Application to Vision Computing, *Kai Tian, Yi Xu, Shuigeng Zhou, Jihong Guan*
213. EV-Gait: Event-Based Robust Gait Recognition Using Dynamic Vision Sensors, *Yanxiang Wang, Bowen Du, Yiran Shen, Kai Wu, Guangrong Zhao, Jianguo Sun, Hongkai Wen*
214. ToothNet: Automatic Tooth Instance Segmentation and Identification From Cone Beam CT Images, *Zhiming Cui, Changjian Li, Wenping Wang*
215. Modularized Textual Grounding for Counterfactual Resilience, *Zhiyuan Fang, Shu Kong, Charless Fowlkes, Yezhou Yang*
216. L3-Net: Towards Learning Based LiDAR Localization for Autonomous Driving, *Weixin Lu, Yao Zhou, Guowei Wan, Shenhua Hou, Shiyu Song*

1130-1330 Doctoral Consortium (Bogarts & Co.) (by invitation only)

Supported by:



- Zhaowei Cai (*Univ. of California San Diego*)
- Jen-Hao Rick Chang (*Carnegie Mellon Univ.*)
- Pelin Dogan (*ETH Zurich*)
- Jiangxin Dong (*Dalian Univ. of Technology*)
- Yueqi Duan (*Tsinghua Univ.*)
- Yang Feng (*Univ. of Rochester*)
- Di Hu (*Northwestern Polytechnical Univ.*)
- Sunghoon Im (*KAIST*)
- Kyungdon Joo (*KAIST*)
- Shu Kong (*Univ. of California, Irvine*)
- Wei-Sheng Lai (*Univ. of California, Merced*)
- Donghoon Lee (*Seoul National Univ.*)
- Yijun Li (*Univ. of California, Merced*)
- Wenbo Li (*Univ. at Albany, State Univ. of New York*)
- Yikang Li (*Chinese Univ. of Hong Kong*)
- Renjie Liao (*Univ. of Toronto*)
- Tsung-Yu Lin (*Univ. of Massachusetts Amherst*)
- Chao Liu (*Carnegie Mellon Univ.*)
- Chen Liu (*Washington Univ. in St. Louis*)
- Davide Moltisanti (*Univ. of Bristol*)
- Seyed-Mohsen Moosavi-Dezfooli (*EPFL*)
- Nils Murrugarra-Llerena (*Univ. of Pittsburgh*)
- Hyeonwoo Noh (*Pohang Univ. of Science and Technology*)
- Filip Radenovic (*Czech Technical Univ. in Prague*)
- Simone Schaub-Meyer (*ETH Zurich*)
- Paul Hongsuck Seo (*Pohang Univ. of Science and Technology*)
- Zhiqiang Shen (*Fudan Univ.*)
- Zheng Shou (*Columbia Univ.*)
- Yumin Suh (*Seoul National Univ.*)
- Lin Sun (*Hong Kong Univ. of Science and Technology*)
- Meng Tang (*Univ. of Waterloo*)
- Luan Tran (*Michigan State Univ.*)
- Gul Varol (*Inria / Ecole Normale Supérieure*)
- Xin Wang (*Univ. of California, Berkeley*)
- Shenlong Wang (*Univ. of Toronto*)
- Hang Zhao (*MIT*)
- Hengshuang Zhao (*Chinese Univ. of Hong Kong*)
- Yipin Zhou (*Univ. of North Carolina at Chapel Hill*)
- Chuhan Zou (*UIUC*)
- Xinxin Zuo (*Univ. of Kentucky*)

1130-1330 Lunch (Pacific Ballroom)

1320–1520 Setup for Poster Session 2-2P (Exhibit Hall)**1330–1520 Oral Session 2-2A: Recognition**
(Terrace Theater)

Papers in this session are in Poster Session 2-2P (Posters 25–42)

Chairs: Abhinav Shrivastava (*Univ. of Maryland*)
Olga Russakovsky (*Princeton Univ.*)Format (5 min. presentation; 3 min. group questions/3 papers)

1. [1330] Panoptic Feature Pyramid Networks, *Alexander Kirillov, Ross Girshick, Kaiming He, Piotr Dollár*
2. [1335] Mask Scoring R-CNN, *Zhaojin Huang, Lichao Huang, Yongchao Gong, Chang Huang, Xinggang Wang*
3. [1340] Reasoning-RCNN: Unifying Adaptive Global Reasoning Into Large-Scale Object Detection, *Hang Xu, Chenhan Jiang, Xiaodan Liang, Liang Lin, Zhenguo Li*
4. [1348] Cross-Modality Personalization for Retrieval, *Nils Murrugarra-Llerena, Adriana Kovashka*
5. [1353] Composing Text and Image for Image Retrieval - an Empirical Odyssey, *Nam Vo, Lu Jiang, Chen Sun, Kevin Murphy, Li-Jia Li, Li Fei-Fei, James Hays*
6. [1358] Arbitrary Shape Scene Text Detection With Adaptive Text Region Representation, *Xiaobing Wang, Yingying Jiang, Zhenbo Luo, Cheng-Lin Liu, Hyunsoo Choi, Sungjin Kim*
7. [1406] Adaptive NMS: Refining Pedestrian Detection in a Crowd, *Songtao Liu, Di Huang, Yunhong Wang*
8. [1411] Point in, Box Out: Beyond Counting Persons in Crowds, *Yuting Liu, Miaoqing Shi, Qijun Zhao, Xiaofang Wang*
9. [1416] Locating Objects Without Bounding Boxes, *Javier Ribera, David Güera, Yuhao Chen, Edward J. Delp*
10. [1424] FineGAN: Unsupervised Hierarchical Disentanglement for Fine-Grained Object Generation and Discovery, *Krishna Kumar Singh, Utkarsh Ojha, Yong Jae Lee*
11. [1429] Mutual Learning of Complementary Networks via Residual Correction for Improving Semi-Supervised Classification, *Si Wu, Jichang Li, Cheng Liu, Zhiwen Yu, Hau-San Wong*
12. [1434] Sampling Techniques for Large-Scale Object Detection From Sparsely Annotated Objects, *Yusuke Niitani, Takuya Akiba, Tommi Kerola, Toru Ogawa, Shotaro Sano, Shuji Suzuki*
13. [1442] Curls & Whey: Boosting Black-Box Adversarial Attacks, *Yucheng Shi, Siyu Wang, Yahong Han*
14. [1447] Barrage of Random Transforms for Adversarially Robust Defense, *Edward Raff, Jared Sylvester, Steven Forsyth, Mark McLean*
15. [1452] Aggregation Cross-Entropy for Sequence Recognition, *Zecheng Xie, Yaoxiong Huang, Yuanzhi Zhu, Lianwen Jin, Yuliang Liu, Lele Xie*
16. [1500] LaSO: Label-Set Operations Networks for Multi-Label Few-Shot Learning, *Amit Alfassy, Leonid Karlinsky, Amit Aides, Joseph Shtok, Sivan Harary, Rogerio Feris, Raja Giryes, Alex M. Bronstein*
17. [1505] Few-Shot Learning With Localization in Realistic Settings, *Davis Wertheimer, Bharath Hariharan*
18. [1510] AdaGraph: Unifying Predictive and Continuous Domain Adaptation Through Graphs, *Massimiliano Mancini, Samuel Rota Bulò, Barbara Caputo, Elisa Ricci*

1330–1520 Oral Session 2-2B: Language & Reasoning
(Grand Ballroom)

Papers in this session are in Poster Session 2-2P (Posters 177–194)

Chairs: Adriana Kovashka (*Univ. of Pittsburgh*)
Yong Jae Lee (*Univ. of California, Davis*)Format (5 min. presentation; 3 min. group questions/3 papers)

1. [1330] Grounded Video Description, *Luowei Zhou, Yannis Kalantidis, Xinlei Chen, Jason J. Corso, Marcus Rohrbach*
2. [1335] Streamlined Dense Video Captioning, *Jonghwan Mun, Linjie Yang, Zhou Ren, Ning Xu, Bohyung Han*
3. [1340] Adversarial Inference for Multi-Sentence Video Description, *Jae Sung Park, Marcus Rohrbach, Trevor Darrell, Anna Rohrbach*
4. [1348] Unified Visual-Semantic Embeddings: Bridging Vision and Language With Structured Meaning Representations, *Hao Wu, Jiayuan Mao, Yufeng Zhang, Yuning Jiang, Lei Li, Weiwei Sun, Wei-Ying Ma*
5. [1353] Learning to Compose Dynamic Tree Structures for Visual Contexts, *Kaihua Tang, Hanwang Zhang, Baoyuan Wu, Wenhan Luo, Wei Liu*
6. [1358] Reinforced Cross-Modal Matching and Self-Supervised Imitation Learning for Vision-Language Navigation, *Xin Wang, Qiuyuan Huang, Asli Celikyilmaz, Jianfeng Gao, Dinghan Shen, Yuan-Fang Wang, William Yang Wang, Lei Zhang*
7. [1406] Dynamic Fusion With Intra- and Inter-Modality Attention Flow for Visual Question Answering, *Peng Gao, Zhengkai Jiang, Haoxuan You, Pan Lu, Steven C. H. Hoi, Xiaogang Wang, Hongsheng Li*
8. [1411] Cycle-Consistency for Robust Visual Question Answering, *Meet Shah, Xinlei Chen, Marcus Rohrbach, Devi Parikh*
9. [1416] Embodied Question Answering in Photorealistic Environments With Point Cloud Perception, *Erik Wijmans, Samyak Datta, Oleksandr Maksymets, Abhishek Das, Georgia Gkioxari, Stefan Lee, Irfan Essa, Devi Parikh, Dhruv Batra*
10. [1424] Reasoning Visual Dialogs With Structural and Partial Observations, *Zilong Zheng, Wenguan Wang, Siyuan Qi, Song-Chun Zhu*
11. [1429] Recursive Visual Attention in Visual Dialog, *Yulei Niu, Hanwang Zhang, Manli Zhang, Jianhong Zhang, Zhiwu Lu, Ji-Rong Wen*
12. [1434] Two Body Problem: Collaborative Visual Task Completion, *Unnat Jain, Luca Weihs, Eric Kolve, Mohammad Rastegari, Svetlana Lazebnik, Ali Farhadi, Alexander G. Schwing, Aniruddha Kembhavi*
13. [1442] GQA: A New Dataset for Real-World Visual Reasoning and Compositional Question Answering, *Drew A. Hudson, Christopher D. Manning*
14. [1447] Text2Scene: Generating Compositional Scenes From Textual Descriptions, *Fuwen Tan, Song Feng, Vicente Ordonez*
15. [1452] From Recognition to Cognition: Visual Commonsense Reasoning, *Rowan Zellers, Yonatan Bisk, Ali Farhadi, Yejin Choi*
16. [1500] The Regretful Agent: Heuristic-Aided Navigation Through Progress Estimation, *Chih-Yao Ma, Zuxuan Wu, Ghassan AlRegib, Caiming Xiong, Zsolt Kira*
17. [1505] Tactical Rewind: Self-Correction via Backtracking in Vision-And-Language Navigation, *Liyiming Ke, Xiujun Li, Yonatan Bisk, Ari Holtzman, Zhe Gan, Jingjing Liu, Jianfeng Gao, Yejin Choi, Siddhartha Srinivasa*

18. [1510] Learning to Learn How to Learn: Self-Adaptive Visual Navigation Using Meta-Learning, *Mitchell Wortsman, Kiana Ehsani, Mohammad Rastegari, Ali Farhadi, Roozbeh Mottaghi*

1330–1520 Oral Session 2-2C: Computational Photography & Graphics
(Promenade Ballroom)

Papers in this session are in Poster Session 2-2P (Posters 130–147)

Chairs: Sanjeev Koppal (*Univ. of Florida*)
Jingyi Yu (*Shanghai Tech Univ.*)

Format (5 min. presentation; 3 min. group questions/3 papers)

1. [1330] Photon-Flooded Single-Photon 3D Cameras, *Anant Gupta, Atul Ingle, Andreas Velten, Mohit Gupta*
2. [1335] High Flux Passive Imaging With Single-Photon Sensors, *Atul Ingle, Andreas Velten, Mohit Gupta*
3. [1340] Acoustic Non-Line-Of-Sight Imaging, *David B. Lindell, Gordon Wetzstein, Vladlen Koltun*
4. [1348] Steady-State Non-Line-Of-Sight Imaging, *Wenzheng Chen, Simon Daneau, Fahim Mannan, Felix Heide*
5. [1353] A Theory of Fermat Paths for Non-Line-Of-Sight Shape Reconstruction, *Shumian Xin, Sotiris Nousias, Kiriakos N. Kutulakos, Aswin C. Sankaranarayanan, Srinivasa G. Narasimhan, Ioannis Gkioulekas*
6. [1358] End-To-End Projector Photometric Compensation, *Bingyao Huang, Haibin Ling*
7. [1406] Bringing a Blurry Frame Alive at High Frame-Rate With an Event Camera, *Liyuan Pan, Cedric Scheerlinck, Xin Yu, Richard Hartley, Miaomiao Liu, Yuchao Dai*
8. [1411] Bringing Alive Blurred Moments, *Kuldeep Purohit, Anshul Shah, A. N. Rajagopalan*
9. [1416] Learning to Synthesize Motion Blur, *Tim Brooks, Jonathan T. Barron*
10. [1424] Underexposed Photo Enhancement Using Deep Illumination Estimation, *Ruixing Wang, Qing Zhang, Chi-Wing Fu, Xiaoyong Shen, Wei-Shi Zheng, Jiaya Jia*
11. [1429] Blind Visual Motif Removal From a Single Image, *Amir Hertz, Sharon Fogel, Rana Hanocka, Raja Giryes, Daniel Cohen-Or*
12. [1434] Non-Local Meets Global: An Integrated Paradigm for Hyperspectral Denoising, *Wei He, Quanming Yao, Chao Li, Naoto Yokoya, Qibin Zhao*
13. [1442] Neural Rerendering in the Wild, *Moustafa Meshry, Dan B. Goldman, Sameh Khamis, Hugues Hoppe, Rohit Pandey, Noah Snavely, Ricardo Martin-Brualla*
14. [1447] GeoNet: Deep Geodesic Networks for Point Cloud Analysis, *Tong He, Haibin Huang, Li Yi, Yuqian Zhou, Chihao Wu, Jue Wang, Stefano Soatto*
15. [1452] MeshAdv: Adversarial Meshes for Visual Recognition, *Chaowei Xiao, Dawei Yang, Bo Li, Jia Deng, Mingyan Liu*
16. [1500] Fast Spatially-Varying Indoor Lighting Estimation, *Mathieu Garon, Kalyan Sunkavalli, Sunil Hadap, Nathan Carr, Jean-François Lalonde*
17. [1505] Neural Illumination: Lighting Prediction for Indoor Environments, *Shuran Song, Thomas Funkhouser*
18. [1510] Deep Sky Modeling for Single Image Outdoor Lighting Estimation, *Yannick Hold-Geoffroy, Akshaya Athawale, Jean-François Lalonde*

1520–1620 Afternoon Break (Exhibit Hall)

1520–1800 Demos (Exhibit Hall)

- Human-In-The-Loop Framework for Land Cover Prediction, *Caleb Robinson, Le Hou, Kolya Malkin, Rachel Soobitsky, Jacob Czawlytko, Bistra Dilkina, Nebojsa Jojic* (*Georgia Institute of Technology*)
- SensitiveNets Demo: Learning Agnostic Representations, *Aythami Morales, Julian Fierrez, Ruben Vera, Ruben Tolosana* (*Universidad Autonoma de Madrid*)
- Towards Extreme Resolution 3D Imaging With Low-Cost Time-Of-Flight Cameras, *Felipe Gutierrez-Barragan, Andreas Velten, Mohit Gupta* (*Univ. of Wisconsin-Madison*)
- End-To-End Pipeline of Document Information Extraction Over Mobile Phones, *Kai Chen* (*Shanghai Jiaotong Univ.*)

1520–1800 Exhibits (Exhibit Hall)

- See Exhibits map for list of exhibitors.

1520–1800 Poster Session 2-2P (Exhibit Hall)

Deep Learning

1. Bidirectional Learning for Domain Adaptation of Semantic Segmentation, *Yunsheng Li, Lu Yuan, Nuno Vasconcelos*
2. Enhanced Bayesian Compression via Deep Reinforcement Learning, *Xin Yuan, Liangliang Ren, Jiwen Lu, Jie Zhou*
3. Strong-Weak Distribution Alignment for Adaptive Object Detection, *Kuniaki Saito, Yoshitaka Ushiku, Tatsuya Harada, Kate Saenko*
4. MFAS: Multimodal Fusion Architecture Search, *Juan-Manuel Pérez-Rúa, Valentin Vielzeuf, Stéphane Pateux, Moez Baccouche, Frederic Jurie*
5. Disentangling Adversarial Robustness and Generalization, *David Stutz, Matthias Hein, Bernt Schiele*
6. ShieldNets: Defending Against Adversarial Attacks Using Probabilistic Adversarial Robustness, *Rajkumar Theagarajan, Ming Chen, Bir Bhanu, Jing Zhang*
7. Deeply-Supervised Knowledge Synergy, *Dawei Sun, Anbang Yao, Aojun Zhou, Hao Zhao*
8. Dual Residual Networks Leveraging the Potential of Paired Operations for Image Restoration, *Xing Liu, Masanori Suganuma, Zhun Sun, Takayuki Okatani*
9. Probabilistic End-To-End Noise Correction for Learning With Noisy Labels, *Kun Yi, Jianxin Wu*
10. Attention-Guided Unified Network for Panoptic Segmentation, *Yanwei Li, Xinze Chen, Zheng Zhu, Lingxi Xie, Guan Huang, Dalong Du, Xingang Wang*
11. NAS-FPN: Learning Scalable Feature Pyramid Architecture for Object Detection, *Golnaz Ghiasi, Tsung-Yi Lin, Quoc V. Le*
12. OICSR: Out-In-Channel Sparsity Regularization for Compact Deep Neural Networks, *Jiashi Li, Qi Qi, Jingyu Wang, Ce Ge, Yujian Li, Zhangzhang Yue, Haifeng Sun*
13. Semantically Aligned Bias Reducing Zero Shot Learning, *Akanksha Paul, Narayanan C. Krishnan, Prateek Munjal*
14. Feature Space Perturbations Yield More Transferable Adversarial Examples, *Nathan Inkawhich, Wei Wen, Hai (Helen) Li, Yiran Chen*

15. IGE-Net: Inverse Graphics Energy Networks for Human Pose Estimation and Single-View Reconstruction, *Dominic Jack, Frederic Maire, Sareh Shirazi, Anders Eriksson*
 16. Accelerating Convolutional Neural Networks via Activation Map Compression, *Georgios Georgiadis*
 17. Knowledge Distillation via Instance Relationship Graph, *Yufan Liu, Jiajiong Cao, Bing Li, Chunfeng Yuan, Weiming Hu, Yangxi Li, Yunqiang Duan*
 18. PPGNet: Learning Point-Pair Graph for Line Segment Detection, *Ziheng Zhang, Zhengxin Li, Ning Bi, Jia Zheng, Jinlei Wang, Kun Huang, Weixin Luo, Yanyu Xu, Shenghua Gao*
 19. Building Detail-Sensitive Semantic Segmentation Networks With Polynomial Pooling, *Zhen Wei, Jingyi Zhang, Li Liu, Fan Zhu, Fumin Shen, Yi Zhou, Si Liu, Yao Sun, Ling Shao*
 20. Variational Bayesian Dropout With a Hierarchical Prior, *Yuhang Liu, Wenyong Dong, Lei Zhang, Dong Gong, Qinfeng Shi*
 21. AANet: Attribute Attention Network for Person Re-Identifications, *Chiat-Pin Tay, Sharmili Roy, Kim-Hui Yap*
 22. Overcoming Limitations of Mixture Density Networks: A Sampling and Fitting Framework for Multimodal Future Prediction, *Osama Makansi, Eddy Ilg, Özgün Çiçek, Thomas Brox*
 23. A Main/Subsidiary Network Framework for Simplifying Binary Neural Networks, *Yinghao Xu, Xin Dong, Yudian Li, Hao Su*
 24. PointNetLK: Robust & Efficient Point Cloud Registration Using PointNet, *Yasuhiro Aoki, Hunter Goforth, Rangaprasad Arun Srivatsan, Simon Lucey*
- Recognition**
25. Panoptic Feature Pyramid Networks, *Alexander Kirillov, Ross Girshick, Kaiming He, Piotr Dollár*
 26. Mask Scoring R-CNN, *Zhaojin Huang, Lichao Huang, Yongchao Gong, Chang Huang, Xinggang Wang*
 27. Reasoning-RCNN: Unifying Adaptive Global Reasoning Into Large-Scale Object Detection, *Hang Xu, Chenhan Jiang, Xiaodan Liang, Liang Lin, Zhenguo Li*
 28. Cross-Modality Personalization for Retrieval, *Nils Murrugarra-Llerena, Adriana Kovashka*
 29. Composing Text and Image for Image Retrieval - an Empirical Odyssey, *Nam Vo, Lu Jiang, Chen Sun, Kevin Murphy, Li-Jia Li, Li Fei-Fei, James Hays*
 30. Arbitrary Shape Scene Text Detection With Adaptive Text Region Representation, *Xiaobing Wang, Yingying Jiang, Zhenbo Luo, Cheng-Lin Liu, Hyunsoo Choi, Sungjin Kim*
 31. Adaptive NMS: Refining Pedestrian Detection in a Crowd, *Songtao Liu, Di Huang, Yunhong Wang*
 32. Point in, Box Out: Beyond Counting Persons in Crowds, *Yuting Liu, Miaoqing Shi, Qijun Zhao, Xiaofang Wang*
 33. Locating Objects Without Bounding Boxes, *Javier Ribera, David Güera, Yuhao Chen, Edward J. Delp*
 34. FineGAN: Unsupervised Hierarchical Disentanglement for Fine-Grained Object Generation and Discovery, *Krishna Kumar Singh, Utkarsh Ojha, Yong Jae Lee*
 35. Mutual Learning of Complementary Networks via Residual Correction for Improving Semi-Supervised Classification, *Si Wu, Jichang Li, Cheng Liu, Zhiwen Yu, Hau-San Wong*
 36. Sampling Techniques for Large-Scale Object Detection From Sparsely Annotated Objects, *Yusuke Niitani, Takuya Akiba, Tommi Kerola, Toru Ogawa, Shotaro Sano, Shuji Suzuki*
 37. Curls & Whey: Boosting Black-Box Adversarial Attacks, *Yucheng Shi, Siyu Wang, Yahong Han*
 38. Barrage of Random Transforms for Adversarially Robust Defense, *Edward Raff, Jared Sylvester, Steven Forsyth, Mark McLean*
 39. Aggregation Cross-Entropy for Sequence Recognition, *Zecheng Xie, Yaoxiong Huang, Yuanzhi Zhu, Lianwen Jin, Yuliang Liu, Lele Xie*
 40. LaSO: Label-Set Operations Networks for Multi-Label Few-Shot Learning, *Amit Alfassy, Leonid Karlinsky, Amit Aides, Joseph Shtok, Sivan Harary, Rogerio Feris, Raja Giryes, Alex M. Bronstein*
 41. Few-Shot Learning With Localization in Realistic Settings, *Davis Wertheimer, Bharath Hariharan*
 42. AdaGraph: Unifying Predictive and Continuous Domain Adaptation Through Graphs, *Massimiliano Mancini, Samuel Rota Bulò, Barbara Caputo, Elisa Ricci*
 43. Few-Shot Adaptive Faster R-CNN, *Tao Wang, Xiaopeng Zhang, Li Yuan, Jiashi Feng*
 44. VRSTC: Occlusion-Free Video Person Re-Identification, *Ruibing Hou, Bingpeng Ma, Hong Chang, Xinqian Gu, Shiguang Shan, Xilin Chen*
 45. Compact Feature Learning for Multi-Domain Image Classification, *Yajing Liu, Xinmei Tian, Ya Li, Zhiwei Xiong, Feng Wu*
 46. Adaptive Transfer Network for Cross-Domain Person Re-Identification, *Jiawei Liu, Zheng-Jun Zha, Di Chen, Richang Hong, Meng Wang*
 47. Large-Scale Few-Shot Learning: Knowledge Transfer With Class Hierarchy, *Aoxue Li, Tiange Luo, Zhiwu Lu, Tao Xiang, Liwei Wang*
 48. Moving Object Detection Under Discontinuous Change in Illumination Using Tensor Low-Rank and Invariant Sparse Decomposition, *Moein Shakeri, Hong Zhang*
 49. Pedestrian Detection With Autoregressive Network Phases, *Garrick Brazil, Xiaoming Liu*
 50. All You Need Is a Few Shifts: Designing Efficient Convolutional Neural Networks for Image Classification, *Weijie Chen, Di Xie, Yuan Zhang, Shiliang Pu*
 51. Stochastic Class-Based Hard Example Mining for Deep Metric Learning, *Yumin Suh, Bohyung Han, Wonsik Kim, Kyoung Mu Lee*
 52. Revisiting Local Descriptor Based Image-To-Class Measure for Few-Shot Learning, *Wenbin Li, Lei Wang, Jinglin Xu, Jing Huo, Yang Gao, Jiebo Luo*
 53. Towards Robust Curve Text Detection With Conditional Spatial Expansion, *Zichuan Liu, Guosheng Lin, Sheng Yang, Fayao Liu, Weisi Lin, Wang Ling Goh*
 54. Revisiting Perspective Information for Efficient Crowd Counting, *Miaoqing Shi, Zhaohui Yang, Chao Xu, Qijun Chen*
 55. Towards Universal Object Detection by Domain Attention, *Xudong Wang, Zhaowei Cai, Dashan Gao, Nuno Vasconcelos*
 56. Ensemble Deep Manifold Similarity Learning Using Hard Proxies, *Nicolas Aziere, Sinisa Todorovic*
 57. Quantization Networks, *Jiwei Yang, Xu Shen, Jun Xing, Xinmei Tian, Houqiang Li, Bing Deng, Jianqiang Huang, Xian-sheng Hua*
 58. RES-PCA: A Scalable Approach to Recovering Low-Rank Matrices, *Chong Peng, Chenglizhao Chen, Zhao Kang, Jianbo Li, Qiang Cheng*
 59. Occlusion-Net: 2D/3D Occluded Keypoint Localization Using Graph Networks, *N. Dinesh Reddy, Minh Vo, Srinivasa G. Narasimhan*

60. Efficient Featurized Image Pyramid Network for Single Shot Detector, *Yanwei Pang, Tiancai Wang, Rao Muhammad Anwer, Fahad Shahbaz Khan, Ling Shao*
61. Multi-Task Multi-Sensor Fusion for 3D Object Detection, *Ming Liang, Bin Yang, Yun Chen, Rui Hu, Raquel Urtasun*
62. Domain-Specific Batch Normalization for Unsupervised Domain Adaptation, *Woong-Gi Chang, Tackgeun You, Seonguk Seo, Suha Kwak, Bohyung Han*
63. Grid R-CNN, *Xin Lu, Buyu Li, Yuxin Yue, Quanguan Li, Junjie Yan*
64. MetaCleaner: Learning to Hallucinate Clean Representations for Noisy-Labeled Visual Recognition, *Weihe Zhang, Yali Wang, Yu Qiao*
65. Mapping, Localization and Path Planning for Image-Based Navigation Using Visual Features and Map, *Janine Thoma, Danda Pani Paudel, Ajad Chhatkuli, Thomas Probst, Luc Van Gool*
66. Triply Supervised Decoder Networks for Joint Detection and Segmentation, *Jiale Cao, Yanwei Pang, Xuelong Li*
67. Leveraging the Invariant Side of Generative Zero-Shot Learning, *Jingjing Li, Mengmeng Jing, Ke Lu, Zhengming Ding, Lei Zhu, Zi Huang*
68. Exploring the Bounds of the Utility of Context for Object Detection, *Ehud Barnea, Ohad Ben-Shahar*

Segmentation, Grouping, & Shape

69. A-CNN: Annularly Convolutional Neural Networks on Point Clouds, *Artem Komarichev, Zichun Zhong, Jing Hua*
70. DARNet: Deep Active Ray Network for Building Segmentation, *Dominic Cheng, Renjie Liao, Sanja Fidler, Raquel Urtasun*
71. Point Cloud Oversegmentation With Graph-Structured Deep Metric Learning, *Loic Landrieu, Mohamed Boussaha*
72. Graphonomy: Universal Human Parsing via Graph Transfer Learning, *Ke Gong, Yiming Gao, Xiaodan Liang, Xiaohui Shen, Meng Wang, Liang Lin*
73. Fitting Multiple Heterogeneous Models by Multi-Class Cascaded T-Linkage, *Luca Magri, Andrea Fusiello*
74. A Late Fusion CNN for Digital Matting, *Yunke Zhang, Lixue Gong, Lubin Fan, Peiran Ren, Qixing Huang, Hujun Bao, Weiwei Xu*
75. BASNet: Boundary-Aware Salient Object Detection, *Xuebin Qin, Zichen Zhang, Chenyang Huang, Chao Gao, Masood Dehghan, Martin Jagersand*
76. ZigZagNet: Fusing Top-Down and Bottom-Up Context for Object Segmentation, *Di Lin, Dingguo Shen, Siting Shen, Yuanfeng Ji, Dani Lischinski, Daniel Cohen-Or, Hui Huang*
77. Object Instance Annotation With Deep Extreme Level Set Evolution, *Zian Wang, David Acuna, Huan Ling, Amlan Kar, Sanja Fidler*
78. Leveraging Crowdsourced GPS Data for Road Extraction From Aerial Imagery, *Tao Sun, Zonglin Di, Pengyu Che, Chun Liu, Yin Wang*
79. Adaptive Pyramid Context Network for Semantic Segmentation, *Junjun He, Zhongying Deng, Lei Zhou, Yali Wang, Yu Qiao*

Statistics, Physics, Theory, & Datasets

80. Isospectralization, or How to Hear Shape, Style, and Correspondence, *Luca Cosmo, Mikhail Panine, Arianna Rampini, Maks Ovsjanikov, Michael M. Bronstein, Emanuele Rodolà*
81. Speech2Face: Learning the Face Behind a Voice, *Tae-Hyun Oh, Tali Dekel, Changil Kim, Inbar Mosseri, William T. Freeman, Michael Rubinstein, Wojciech Matusik*

82. Joint Manifold Diffusion for Combining Predictions on Decoupled Observations, *Kwang In Kim, Hyung Jin Chang*
83. Audio Visual Scene-Aware Dialog, *Huda Alamri, Vincent Cartillier, Abhishek Das, Jue Wang, Anoop Cherian, Irfan Essa, Dhruv Batra, Tim K. Marks, Chiori Hori, Peter Anderson, Stefan Lee, Devi Parikh*
84. Learning to Minify Photometric Stereo, *Junxuan Li, Antonio Robles-Kelly, Shaodi You, Yasuyuki Matsushita*
85. Reflective and Fluorescent Separation Under Narrow-Band Illumination, *Koji Koyamatsu, Daichi Hidaka, Takahiro Okabe, Hendrik P. A. Lensch*
86. Depth From a Polarisation + RGB Stereo Pair, *Dizhong Zhu, William A. P. Smith*
87. Rethinking the Evaluation of Video Summaries, *Mayu Otani, Yuta Nakashima, Esa Rahtu, Janne Heikkilä*
88. What Object Should I Use? - Task Driven Object Detection, *Johann Sawatzky, Yaser Souri, Christian Grund, Jürgen Gall*

3D Multiview

89. Triangulation Learning Network: From Monocular to Stereo 3D Object Detection, *Zengyi Qin, Jinglu Wang, Yan Lu*
90. Connecting the Dots: Learning Representations for Active Monocular Depth Estimation, *Gernot Riegler, Yiyi Liao, Simon Donné, Vladlen Koltun, Andreas Geiger*
91. Learning Non-Volumetric Depth Fusion Using Successive Reprojections, *Simon Donné, Andreas Geiger*
92. Stereo R-CNN Based 3D Object Detection for Autonomous Driving, *Peiliang Li, Xiaozhi Chen, Shaojie Shen*
93. Hybrid Scene Compression for Visual Localization, *Federico Camposeco, Andrea Cohen, Marc Pollefeys, Torsten Sattler*

3D Single View & RGBD

94. MMFace: A Multi-Metric Regression Network for Unconstrained Face Reconstruction, *Hongwei Yi, Chen Li, Qiong Cao, Xiaoyong Shen, Sheng Li, Guoping Wang, Yu-Wing Tai*
95. 3D Motion Decomposition for RGBD Future Dynamic Scene Synthesis, *Xiaojuan Qi, Zhengzhe Liu, Qifeng Chen, Jiaya Jia*
96. Single Image Depth Estimation Trained via Depth From Defocus Cues, *Shir Gur, Lior Wolf*
97. RGBD Based Dimensional Decomposition Residual Network for 3D Semantic Scene Completion, *Jie Li, Yu Liu, Dong Gong, Qinfeng Shi, Xia Yuan, Chunxia Zhao, Ian Reid*
98. Neural Scene Decomposition for Multi-Person Motion Capture, *Helge Rhodin, Victor Constantin, Isinsu Katircioglu, Mathieu Salzmann, Pascal Fua*

Face & Body

99. Efficient Decision-Based Black-Box Adversarial Attacks on Face Recognition, *Yinpeng Dong, Hang Su, Baoyuan Wu, Zhifeng Li, Wei Liu, Tong Zhang, Jun Zhu*
100. FA-RPN: Floating Region Proposals for Face Detection, *Mahyar Najibi, Bharat Singh, Larry S. Davis*
101. Bayesian Hierarchical Dynamic Model for Human Action Recognition, *Rui Zhao, Wanru Xu, Hui Su, Qiang Ji*
102. Mixed Effects Neural Networks (MeNets) With Applications to Gaze Estimation, *Yunyang Xiong, Hyunwoo J. Kim, Vikas Singh*
103. 3D Human Pose Estimation in Video With Temporal Convolutions and Semi-Supervised Training, *Dario Pavllo, Christoph Feichtenhofer, David Grangier, Michael Auli*

104. Learning to Regress 3D Face Shape and Expression From an Image Without 3D Supervision, *Soubhik Sanyal, Timo Bolkart, Haiwen Feng, Michael J. Black*
105. PoseFix: Model-Agnostic General Human Pose Refinement Network, *Gyeongsik Moon, Ju Yong Chang, Kyoung Mu Lee*
106. RepNet: Weakly Supervised Training of an Adversarial Reprojection Network for 3D Human Pose Estimation, *Bastian Wandt, Bodo Rosenhahn*
107. Fast and Robust Multi-Person 3D Pose Estimation From Multiple Views, *Junting Dong, Wen Jiang, Qixing Huang, Hujun Bao, Xiaowei Zhou*
108. Face-Focused Cross-Stream Network for Deception Detection in Videos, *Mingyu Ding, An Zhao, Zhiwu Lu, Tao Xiang, Ji-Rong Wen*
109. Unequal-Training for Deep Face Recognition With Long-Tailed Noisy Data, *Yaoyao Zhong, Weihong Deng, Mei Wang, Jiani Hu, Jianteng Peng, Xunqiang Tao, Yaohai Huang*
110. T-Net: Parametrizing Fully Convolutional Nets With a Single High-Order Tensor, *Jean Kossaifi, Adrian Bulat, Georgios Tzimiropoulos, Maja Pantic*
111. Hierarchical Cross-Modal Talking Face Generation With Dynamic Pixel-Wise Loss, *Lele Chen, Ross K. Maddox, Zhiyao Duan, Chenliang Xu*

Action & Video

112. Object-Centric Auto-Encoders and Dummy Anomalies for Abnormal Event Detection in Video, *Radu Tudor Ionescu, Fahad Shahbaz Khan, Mariana-Iuliana Georgescu, Ling Shao*
113. DDLSTM: Dual-Domain LSTM for Cross-Dataset Action Recognition, *Toby Perrett, Dima Damen*
114. The Pros and Cons: Rank-Aware Temporal Attention for Skill Determination in Long Videos, *Hazel Doughty, Walterio Mayol-Cuevas, Dima Damen*
115. Collaborative Spatiotemporal Feature Learning for Video Action Recognition, *Chao Li, Qiaoyong Zhong, Di Xie, Shiliang Pu*
116. MARS: Motion-Augmented RGB Stream for Action Recognition, *Nieves Crasto, Philippe Weinzaepfel, Karteek Alahari, Cordelia Schmid*
117. Convolutional Relational Machine for Group Activity Recognition, *Sina Mokhtarzadeh Azar, Mina Ghadimi Atigh, Ahmad Nickabadi, Alexandre Alahi*
118. Video Summarization by Learning From Unpaired Data, *Mrigank Rochan, Yang Wang*
119. Skeleton-Based Action Recognition With Directed Graph Neural Networks, *Lei Shi, Yifan Zhang, Jian Cheng, Hanqing Lu*
120. PA3D: Pose-Action 3D Machine for Video Recognition, *An Yan, Yali Wang, Zhifeng Li, Yu Qiao*
121. Deep Dual Relation Modeling for Egocentric Interaction Recognition, *Haoxin Li, Yijun Cai, Wei-Shi Zheng*

Motion & Biometrics

122. MOTS: Multi-Object Tracking and Segmentation, *Paul Voigtlaender, Michael Krause, Aljosa Osep, Jonathon Luiten, Berin Balachandar Gnana Sekar, Andreas Geiger, Bastian Leibe*
123. Siamese Cascaded Region Proposal Networks for Real-Time Visual Tracking, *Heng Fan, Haibin Ling*
124. PointFlowNet: Learning Representations for Rigid Motion Estimation From Point Clouds, *Aseem Behl, Despoina Paschalidou, Simon Donn e, Andreas Geiger*

Synthesis

125. Listen to the Image, *Di Hu, Dong Wang, Xuelong Li, Feiping Nie, Qi Wang*
126. Image Super-Resolution by Neural Texture Transfer, *Zhifei Zhang, Zhaowen Wang, Zhe Lin, Hairong Qi*
127. Conditional Adversarial Generative Flow for Controllable Image Synthesis, *Rui Liu, Yu Liu, Xinyu Gong, Xiaogang Wang, Hongsheng Li*
128. How to Make a Pizza: Learning a Compositional Layer-Based GAN Model, *Dim P. Papadopoulos, Youssef Tamaazousti, Ferda Ofli, Ingmar Weber, Antonio Torralba*
129. TransGaGa: Geometry-Aware Unsupervised Image-To-Image Translation, *Wayne Wu, Kaidi Cao, Cheng Li, Chen Qian, Chen Change Loy*

Computational Photography & Graphics

130. Photon-Flooded Single-Photon 3D Cameras, *Anant Gupta, Atul Ingle, Andreas Velten, Mohit Gupta*
131. High Flux Passive Imaging With Single-Photon Sensors, *Atul Ingle, Andreas Velten, Mohit Gupta*
132. Acoustic Non-Line-Of-Sight Imaging, *David B. Lindell, Gordon Wetzstein, Vladlen Koltun*
133. Steady-State Non-Line-Of-Sight Imaging, *Wenzheng Chen, Simon Daneau, Fahim Mannan, Felix Heide*
134. A Theory of Fermat Paths for Non-Line-Of-Sight Shape Reconstruction, *Shumian Xin, Sotiris Nouisias, Kiriakos N. Kutulakos, Aswin C. Sankaranarayanan, Srinivasa G. Narasimhan, Ioannis Gkioulekas*
135. End-To-End Projector Photometric Compensation, *Bingyao Huang, Haibin Ling*
136. Bringing a Blurry Frame Alive at High Frame-Rate With an Event Camera, *Liyuan Pan, Cedric Scheerlinck, Xin Yu, Richard Hartley, Miaomiao Liu, Yuchao Dai*
137. Bringing Alive Blurred Moments, *Kuldeep Purohit, Anshul Shah, A. N. Rajagopalan*
138. Learning to Synthesize Motion Blur, *Tim Brooks, Jonathan T. Barron*
139. Underexposed Photo Enhancement Using Deep Illumination Estimation, *Ruixing Wang, Qing Zhang, Chi-Wing Fu, Xiaoyong Shen, Wei-Shi Zheng, Jiaya Jia*
140. Blind Visual Motif Removal From a Single Image, *Amir Hertz, Sharon Fogel, Rana Hanocka, Raja Giryes, Daniel Cohen-Or*
141. Non-Local Meets Global: An Integrated Paradigm for Hyperspectral Denoising, *Wei He, Quanming Yao, Chao Li, Naoto Yokoya, Qibin Zhao*
142. Neural Rerendering in the Wild, *Moustafa Meshry, Dan B. Goldman, Sameh Khamis, Hugues Hoppe, Rohit Pandey, Noah Snavely, Ricardo Martin-Brualla*
143. GeoNet: Deep Geodesic Networks for Point Cloud Analysis, *Tong He, Haibin Huang, Li Yi, Yuqian Zhou, Chihao Wu, Jue Wang, Stefano Soatto*
144. MeshAdv: Adversarial Meshes for Visual Recognition, *Chaowei Xiao, Dawei Yang, Bo Li, Jia Deng, Mingyan Liu*
145. Fast Spatially-Varying Indoor Lighting Estimation, *Mathieu Garon, Kalyan Sunkavalli, Sunil Hadap, Nathan Carr, Jean-Fran ois Lalonde*
146. Neural Illumination: Lighting Prediction for Indoor Environments, *Shuran Song, Thomas Funkhouser*

- 147. Deep Sky Modeling for Single Image Outdoor Lighting Estimation, *Yannick Hold-Geoffroy, Akshaya Athawale, Jean-François Lalonde*
- 148. Depth-Attentional Features for Single-Image Rain Removal, *Xiaowei Hu, Chi-Wing Fu, Lei Zhu, Pheng-Ann Heng*
- 149. Hyperspectral Image Reconstruction Using a Deep Spatial-Spectral Prior, *Lizhi Wang, Chen Sun, Ying Fu, Min H. Kim, Hua Huang*
- 150. LiFF: Light Field Features in Scale and Depth, *Donald G. Dansereau, Bernd Girod, Gordon Wetzstein*
- 151. Deep Exemplar-Based Video Colorization, *Bo Zhang, Mingming He, Jing Liao, Pedro V. Sander, Lu Yuan, Amine Bermak, Dong Chen*
- 152. On Finding Gray Pixels, *Yanlin Qian, Joni-Kristian Kämäräinen, Jarno Nikkanen, Jiří Matas*

Low-Level & Optimization

- 153. UnOS: Unified Unsupervised Optical-Flow and Stereo-Depth Estimation by Watching Videos, *Yang Wang, Peng Wang, Zhenheng Yang, Chenxu Luo, Yi Yang, Wei Xu*
- 154. Learning Transformation Synchronization, *Xiangru Huang, Zhenxiao Liang, Xiaowei Zhou, Yao Xie, Leonidas J. Guibas, Qixing Huang*
- 155. D2-Net: A Trainable CNN for Joint Description and Detection of Local Features, *Mihai Dusmanu, Ignacio Rocco, Tomas Pajdla, Marc Pollefeys, Josef Sivic, Akihiko Torii, Torsten Sattler*
- 156. Recurrent Neural Networks With Intra-Frame Iterations for Video Deblurring, *Seungjun Nah, Sanghyun Son, Kyoung Mu Lee*
- 157. Learning to Extract Flawless Slow Motion From Blurry Videos, *Meiguang Jin, Zhe Hu, Paolo Favaro*
- 158. Natural and Realistic Single Image Super-Resolution With Explicit Natural Manifold Discrimination, *Jae Woong Soh, Gu Yong Park, Junho Jo, Nam Ik Cho*
- 159. RF-Net: An End-To-End Image Matching Network Based on Receptive Field, *Xuelun Shen, Cheng Wang, Xin Li, Zenglei Yu, Jonathan Li, Chenglu Wen, Ming Cheng, Zijian He*
- 160. Fast Single Image Reflection Suppression via Convex Optimization, *Yang Yang, Wenye Ma, Yin Zheng, Jian-Feng Cai, Weiyu Xu*
- 161. A Mutual Learning Method for Salient Object Detection With Intertwined Multi-Supervision, *Runmin Wu, Mengyang Feng, Wenlong Guan, Dong Wang, Huchuan Lu, Errui Ding*
- 162. Enhanced Pix2pix Dehazing Network, *Yanyun Qu, Yizi Chen, Jingying Huang, Yuan Xie*
- 163. Assessing Personally Perceived Image Quality via Image Features and Collaborative Filtering, *Jari Korhonen*
- 164. Single Image Reflection Removal Exploiting Misaligned Training Data and Network Enhancements, *Kaixuan Wei, Jiaolong Yang, Ying Fu, David Wipf, Hua Huang*

Scenes & Representation

- 165. Exploring Context and Visual Pattern of Relationship for Scene Graph Generation, *Wenbin Wang, Ruiping Wang, Shiguang Shan, Xilin Chen*
- 166. Learning From Synthetic Data for Crowd Counting in the Wild, *Qi Wang, Junyu Gao, Wei Lin, Yuan Yuan*
- 167. A Local Block Coordinate Descent Algorithm for the CSC Model, *Ev Zisselman, Jeremias Sulam, Michael Elad*
- 168. Not Using the Car to See the Sidewalk — Quantifying and Controlling the Effects of Context in Classification and Segmentation, *Rakshith Shetty, Bernt Schiele, Mario Fritz*

- 169. Discovering Fair Representations in the Data Domain, *Novi Quadrianto, Viktoriia Sharmanska, Oliver Thomas*
- 170. Actor-Critic Instance Segmentation, *Nikita Araslanov, Constantin A. Rothkopf, Stefan Roth*
- 171. Generalized Zero- and Few-Shot Learning via Aligned Variational Autoencoders, *Edgar Schönfeld, Sayna Ebrahimi, Samarth Sinha, Trevor Darrell, Zeynep Akata*
- 172. Semantic Projection Network for Zero- and Few-Label Semantic Segmentation, *Yongqin Xian, Subhabrata Choudhury, Yang He, Bernt Schiele, Zeynep Akata*
- 173. GCAN: Graph Convolutional Adversarial Network for Unsupervised Domain Adaptation, *Xinhong Ma, Tianzhu Zhang, Changsheng Xu*
- 174. Seamless Scene Segmentation, *Lorenzo Porzi, Samuel Rota Bulò, Aleksander Colovic, Peter Kotschieder*
- 175. Unsupervised Image Matching and Object Discovery as Optimization, *Huy V. Vo, Francis Bach, Minsu Cho, Kai Han, Yann LeCun, Patrick Pérez, Jean Ponce*
- 176. Wide-Area Crowd Counting via Ground-Plane Density Maps and Multi-View Fusion CNNs, *Qi Zhang, Antoni B. Chan*

Language & Reasoning

- 177. Grounded Video Description, *Luowei Zhou, Yannis Kalantidis, Xinlei Chen, Jason J. Corso, Marcus Rohrbach*
- 178. Streamlined Dense Video Captioning, *Jonghwan Mun, Linjie Yang, Zhou Ren, Ning Xu, Bohyung Han*
- 179. Adversarial Inference for Multi-Sentence Video Description, *Jae Sung Park, Marcus Rohrbach, Trevor Darrell, Anna Rohrbach*
- 180. Unified Visual-Semantic Embeddings: Bridging Vision and Language With Structured Meaning Representations, *Hao Wu, Jiayuan Mao, Yufeng Zhang, Yuning Jiang, Lei Li, Weiwei Sun, Wei-Ying Ma*
- 181. Learning to Compose Dynamic Tree Structures for Visual Contexts, *Kaihua Tang, Hanwang Zhang, Baoyuan Wu, Wenhan Luo, Wei Liu*
- 182. Reinforced Cross-Modal Matching and Self-Supervised Imitation Learning for Vision-Language Navigation, *Xin Wang, Qiuyuan Huang, Asli Celikyilmaz, Jianfeng Gao, Dinghan Shen, Yuan-Fang Wang, William Yang Wang, Lei Zhang*
- 183. Dynamic Fusion With Intra- and Inter-Modality Attention Flow for Visual Question Answering, *Peng Gao, Zhengkai Jiang, Haoxuan You, Pan Lu, Steven C. H. Hoi, Xiaogang Wang, Hongsheng Li*
- 184. Cycle-Consistency for Robust Visual Question Answering, *Meet Shah, Xinlei Chen, Marcus Rohrbach, Devi Parikh*
- 185. Embodied Question Answering in Photorealistic Environments With Point Cloud Perception, *Erik Wijmans, Samyak Datta, Oleksandr Maksymets, Abhishek Das, Georgia Gkioxari, Stefan Lee, Irfan Essa, Devi Parikh, Dhruv Batra*
- 186. Reasoning Visual Dialogs With Structural and Partial Observations, *Zilong Zheng, Wenguan Wang, Siyuan Qi, Song-Chun Zhu*
- 187. Recursive Visual Attention in Visual Dialog, *Yulei Niu, Hanwang Zhang, Manli Zhang, Jianhong Zhang, Zhiwu Lu, Ji-Rong Wen*
- 188. Two Body Problem: Collaborative Visual Task Completion, *Unnat Jain, Luca Weihs, Eric Kolve, Mohammad Rastegari, Svetlana Lazebnik, Ali Farhadi, Alexander G. Schwing, Aniruddha Kembhavi*

Thursday, June 20

0730–1600 Registration (Promenade Atrium & Plaza)

0730–0900 Breakfast (Pacific Ballroom)

0800–1000 Setup for Poster Session 3-1P (Exhibit Hall)

0830–1000 Oral Session 3-1A: Applications
(Terrace Theater)

Papers in this session are in Poster Session 3-1P (Posters 190–204)

Chairs: Yin Li (*Univ. of Wisconsin-Madison*)
Haibin Lin (*Temple Univ.*)

Format (5 min. presentation; 3 min. group questions/3 papers)

1. [0830] Holistic and Comprehensive Annotation of Clinically Significant Findings on Diverse CT Images: Learning From Radiology Reports and Label Ontology, *Ke Yan, Yifan Peng, Veit Sandfort, Mohammadhadi Bagheri, Zhiyong Lu, Ronald M. Summers*
2. [0835] Robust Histopathology Image Analysis: To Label or to Synthesize? *Le Hou, Ayush Agarwal, Dimitris Samaras, Tahsin M. Kurc, Rajarsi R. Gupta, Joel H. Saltz*
3. [0840] Data Augmentation Using Learned Transformations for One-Shot Medical Image Segmentation, *Amy Zhao, Guha Balakrishnan, Frédo Durand, John V. Guttag, Adrian V. Dalca*
4. [0848] Shifting More Attention to Video Salient Object Detection, *Deng-Ping Fan, Wenguan Wang, Ming-Ming Cheng, Jianbing Shen*
5. [0853] Neural Task Graphs: Generalizing to Unseen Tasks From a Single Video Demonstration, *De-An Huang, Suraj Nair, Danfei Xu, Yuke Zhu, Animesh Garg, Li Fei-Fei, Silvio Savarese, Juan Carlos Niebles*
6. [0858] Beyond Tracking: Selecting Memory and Refining Poses for Deep Visual Odometry, *Fei Xue, Xin Wang, Shunkai Li, Qiuyuan Wang, Junqiu Wang, Hongbin Zha*
7. [0906] Image Generation From Layout, *Bo Zhao, Lili Meng, Weidong Yin, Leonid Sigal*
8. [0911] Multimodal Explanations by Predicting Counterfactuality in Videos, *Atsushi Kanehira, Kentaro Takemoto, Sho Inayoshi, Tatsuya Harada*
9. [0916] Learning to Explain With Complemental Examples, *Atsushi Kanehira, Tatsuya Harada*
10. [0924] HAQ: Hardware-Aware Automated Quantization With Mixed Precision, *Kuan Wang, Zhijian Liu, Yujun Lin, Ji Lin, Song Han*
11. [0929] Content Authentication for Neural Imaging Pipelines: End-To-End Optimization of Photo Provenance in Complex Distribution Channels, *Pawel Korus, Nasir Memon*
12. [0934] Inverse Procedural Modeling of Knitwear, *Elena Trunz, Sebastian Merzbach, Jonathan Klein, Thomas Schulze, Michael Weinmann, Reinhard Klein*
13. [0942] Estimating 3D Motion and Forces of Person-Object Interactions From Monocular Video, *Zongmian Li, Jiri Sedlar, Justin Carpentier, Ivan Laptev, Nicolas Mansard, Josef Sivic*
14. [0947] DeepMapping: Unsupervised Map Estimation From Multiple Point Clouds, *Li Ding, Chen Feng*

15. [0952] End-To-End Interpretable Neural Motion Planner, *Wenyuan Zeng, Wenjie Luo, Simon Suo, Abbas Sadat, Bin Yang, Sergio Casas, Raquel Urtasun*

0830–1000 Oral Session 3-1B: Learning, Physics, Theory, & Datasets (Grand Ballroom)

Papers in this session are in Poster Session 3-1P (Posters 77–91)

Chairs: Stephen Gould (*Australian National Univ.*)
Cornelia Fermuller (*Univ. of Maryland, College Park*)

Format (5 min. presentation; 3 min. group questions/3 papers)

1. [0830] Divergence Triangle for Joint Training of Generator Model, Energy-Based Model, and Inferential Model, *Tian Han, Erik Nijkamp, Xiaolin Fang, Mitch Hill, Song-Chun Zhu, Ying Nian Wu*
2. [0835] Image Deformation Meta-Networks for One-Shot Learning, *Zitian Chen, Yanwei Fu, Yu-Xiong Wang, Lin Ma, Wei Liu, Martial Hebert*
3. [0840] Online High Rank Matrix Completion, *Jicong Fan, Madeleine Udell*
4. [0848] Multispectral Imaging for Fine-Grained Recognition of Powders on Complex Backgrounds, *Tiancheng Zhi, Bernardo R. Pires, Martial Hebert, Srinivasa G. Narasimhan*
5. [0853] ContactDB: Analyzing and Predicting Grasp Contact via Thermal Imaging, *Samarth Brahmabhatt, Cusuh Ham, Charles C. Kemp, James Hays*
6. [0858] Robust Subspace Clustering With Independent and Piecewise Identically Distributed Noise Modeling, *Yuanman Li, Jiantao Zhou, Xianwei Zheng, Jinyu Tian, Yuan Yan Tang*
7. [0906] What Correspondences Reveal About Unknown Camera and Motion Models? *Thomas Probst, Ajad Chhatkuli, Danda Pani Paudel, Luc Van Gool*
8. [0911] Self-Calibrating Deep Photometric Stereo Networks, *Guanying Chen, Kai Han, Boxin Shi, Yasuyuki Matsushita, Kwan-Yee K. Wong*
9. [0916] Argoverse: 3D Tracking and Forecasting With Rich Maps, *Ming-Fang Chang, John Lambert, Patsorn Sangkloy, Jagjeet Singh, Slawomir Bak, Andrew Hartnett, De Wang, Peter Carr, Simon Lucey, Deva Ramanan, James Hays*
10. [0924] Side Window Filtering, *Hui Yin, Yuanhao Gong, Guoping Qiu*
11. [0929] Defense Against Adversarial Images Using Web-Scale Nearest-Neighbor Search, *Abhimanyu Dubey, Laurens van der Maaten, Zeki Yalniz, Yixuan Li, Dhruv Mahajan*
12. [0934] Incremental Object Learning From Contiguous Views, *Stefan Stojanov, Samarth Mishra, Ngoc Anh Thai, Nikhil Dhanda, Ahmad Humayun, Chen Yu, Linda B. Smith, James M. Rehg*
13. [0942] IP102: A Large-Scale Benchmark Dataset for Insect Pest Recognition, *Xiaoping Wu, Chi Zhan, Yu-Kun Lai, Ming-Ming Cheng, Jufeng Yang*
14. [0947] CityFlow: A City-Scale Benchmark for Multi-Target Multi-Camera Vehicle Tracking and Re-Identification, *Zheng Tang, Milind Naphade, Ming-Yu Liu, Xiaodong Yang, Stan Birchfield, Shuo Wang, Ratnesh Kumar, David Anastasiu, Jenq-Neng Hwang*
15. [0952] Social-IQ: A Question Answering Benchmark for Artificial Social Intelligence, *Amir Zadeh, Michael Chan, Paul Pu Liang, Edmund Tong, Louis-Philippe Morency*

0830-1000 Oral Session 3-1C: Segmentation & Grouping (Promenade Ballroom)

Papers in this session are in Poster Session 3-1P (Posters 55-69)

Chairs: Stella Yu (*Univ. of California, Berkeley; ICSI*)
Georgia Gkioxari (*Facebook*)

Format (5 min. presentation; 3 min. group questions/3 papers)

1. [0830] UPSNet: A Unified Panoptic Segmentation Network, Yuwen Xiong, Renjie Liao, Hengshuang Zhao, Rui Hu, Min Bai, Ersin Yumer, Raquel Urtasun
2. [0835] JSIS3D: Joint Semantic-Instance Segmentation of 3D Point Clouds With Multi-Task Pointwise Networks and Multi-Value Conditional Random Fields, Quang-Hieu Pham, Thanh Nguyen, Binh-Son Hua, Gemma Roig, Sai-Kit Yeung
3. [0840] Instance Segmentation by Jointly Optimizing Spatial Embeddings and Clustering Bandwidth, Davy Neven, Bert De Brabandere, Marc Proesmans, Luc Van Gool

4. [0848] DeepCO3: Deep Instance Co-Segmentation by Co-Peak Search and Co-Saliency Detection, Kuang-Jui Hsu, Yen-Yu Lin, Yung-Yu Chuang
5. [0853] Improving Semantic Segmentation via Video Propagation and Label Relaxation, Yi Zhu, Karan Sapra, Fitsum A. Reda, Kevin J. Shih, Shawn Newsam, Andrew Tao, Bryan Catanzaro
6. [0858] Accel: A Corrective Fusion Network for Efficient Semantic Segmentation on Video, Samvit Jain, Xin Wang, Joseph E. Gonzalez

7. [0906] Shape2Motion: Joint Analysis of Motion Parts and Attributes From 3D Shapes, Xiaogang Wang, Bin Zhou, Yahao Shi, Xiaowu Chen, Qinqing Zhao, Kai Xu
8. [0911] Semantic Correlation Promoted Shape-Variant Context for Segmentation, Henghui Ding, Xudong Jiang, Bing Shuai, Ai Qun Liu, Gang Wang
9. [0916] Relation-Shape Convolutional Neural Network for Point Cloud Analysis, Yongcheng Liu, Bin Fan, Shiming Xiang, Chunhong Pan

10. [0924] Enhancing Diversity of Defocus Blur Detectors via Cross-Ensemble Network, Wenda Zhao, Bowen Zheng, Qihua Lin, Huchuan Lu
11. [0929] BubbleNets: Learning to Select the Guidance Frame in Video Object Segmentation by Deep Sorting Frames, Brent A. Griffen, Jason J. Corso
12. [0934] Collaborative Global-Local Networks for Memory-Efficient Segmentation of Ultra-High Resolution Images, Wuyang Chen, Ziyu Jiang, Zhangyang Wang, Kexin Cui, Xiaoning Qian

13. [0942] Efficient Parameter-Free Clustering Using First Neighbor Relations, Saquib Sarfraz, Vivek Sharma, Rainer Stiefelhagen
14. [0947] Learning Personalized Modular Network Guided by Structured Knowledge, Xiaodan Liang
15. [0952] A Generative Appearance Model for End-To-End Video Object Segmentation, Joakim Johnander, Martin Danelljan, Emil Brissman, Fahad Shahbaz Khan, Michael Felsberg

1000-1100 Morning Break (Exhibit Hall)

1000-1245 Demos (Exhibit Hall)

- Real-Time Semantic Segmentation Demo Using CFNet, Hang Zhang, Han Zhang, Chenguang Wang, Junyuan Xie (*Amazon Web Services*)

- Demonstration of BioTouchPass: Handwritten Passwords for Touchscreen Biometrics, Ruben Tolosana, Ruben Vera-Rodriguez and Julian Fierrez (*BiDA Lab - Universidad Autonoma de Madrid*)
- Events-To-Video: Real-Time Image Reconstruction With an Event Camera, Henri Rebecq (*Univ. of Zürich*)
- Simulating Circuits From Images, Arthur Chau (*Federal Univ. of Rio de Janeiro*)

1000-1245 Exhibits (Exhibit Hall)

- See Exhibits map for list of exhibitors.

1000-1245 Poster Session 3-1P (Exhibit Hall)

Deep Learning

1. A Flexible Convolutional Solver for Fast Style Transfers, Gilles Puy, Patrick Pérez
2. Cross Domain Model Compression by Structurally Weight Sharing, Shangqian Gao, Cheng Deng, Heng Huang
3. TraVeLGAN: Image-To-Image Translation by Transformation Vector Learning, Matthew Amodio, Smita Krishnaswamy
4. Deep Robust Subjective Visual Property Prediction in Crowdsourcing, Qianqian Xu, Zhiyong Yang, Yangbangyan Jiang, Xiaochun Cao, Qingming Huang, Yuan Yao
5. Transferable AutoML by Model Sharing Over Grouped Datasets, Chao Xue, Junchi Yan, Rong Yan, Stephen M. Chu, Yonggang Hu, Yonghua Lin
6. Learning Not to Learn: Training Deep Neural Networks With Biased Data, Byungju Kim, Hyunwoo Kim, Kyungsu Kim, Sungjin Kim, Junmo Kim
7. IRLAS: Inverse Reinforcement Learning for Architecture Search, Minghao Guo, Zhao Zhong, Wei Wu, Dahua Lin, Junjie Yan
8. Learning for Single-Shot Confidence Calibration in Deep Neural Networks Through Stochastic Inferences, Seonguk Seo, Paul Hongsuck Seo, Bohyung Han
9. Attention-Based Adaptive Selection of Operations for Image Restoration in the Presence of Unknown Combined Distortions, Masanori Suganuma, Xing Liu, Takayuki Okatani
10. Fully Learnable Group Convolution for Acceleration of Deep Neural Networks, Xijun Wang, Meina Kan, Shiguang Shan, Xilin Chen
11. EIGEN: Ecologically-Inspired GENetic Approach for Neural Network Structure Searching From Scratch, Jian Ren, Zhe Li, Jianchao Yang, Ning Xu, Tianbao Yang, David J. Foran
12. Deep Incremental Hashing Network for Efficient Image Retrieval, Dayan Wu, Qi Dai, Jing Liu, Bo Li, Weiping Wang
13. Robustness via Curvature Regularization, and Vice Versa, Seyed-Mohsen Moosavi-Dezfooli, Alhussein Fawzi, Jonathan Uesato, Pascal Frossard
14. SparseFool: A Few Pixels Make a Big Difference, Apostolos Modas, Seyed-Mohsen Moosavi-Dezfooli, Pascal Frossard
15. Interpretable and Fine-Grained Visual Explanations for Convolutional Neural Networks, Jörg Wagner, Jan Mathias Köhler, Tobias Gindele, Leon Hetzel, Jakob Thaddäus Wiedemer, Sven Behnke
16. Structured Pruning of Neural Networks With Budget-Aware Regularization, Carl Lemaire, Andrew Achkar, Pierre-Marc Jodoin
17. MBS: Macroblock Scaling for CNN Model Reduction, Yu-Hsun Lin, Chun-Nan Chou, Edward Y. Chang

18. Fast Neural Architecture Search of Compact Semantic Segmentation Models via Auxiliary Cells, *Vladimir Nekrasov, Hao Chen, Chunhua Shen, Ian Reid*
19. Generating 3D Adversarial Point Clouds, *Chong Xiang, Charles R. Qi, Bo Li*
20. Partial Order Pruning: For Best Speed/Accuracy Trade-Off in Neural Architecture Search, *Xin Li, Yiming Zhou, Zheng Pan, Jiashi Feng*
21. Memory in Memory: A Predictive Neural Network for Learning Higher-Order Non-Stationarity From Spatiotemporal Dynamics, *Yunbo Wang, Jianjin Zhang, Hongyu Zhu, Mingsheng Long, Jianmin Wang, Philip S. Yu*
22. Variational Information Distillation for Knowledge Transfer, *Sungsoo Ahn, Shell Xu Hu, Andreas Damianou, Neil D. Lawrence, Zhenwen Dai*
23. You Look Twice: GaterNet for Dynamic Filter Selection in CNNs, *Zhourong Chen, Yang Li, Samy Bengio, Si Si*
24. SpherePHD: Applying CNNs on a Spherical PolyHeDron Representation of 360° Images, *Yeonkun Lee, Jaeseok Jeong, Jongseob Yun, Wonjune Cho, Kuk-Jin Yoon*
25. ESPNetv2: A Light-Weight, Power Efficient, and General Purpose Convolutional Neural Network, *Sachin Mehta, Mohammad Rastegari, Linda Shapiro, Hannaneh Hajishirzi*
26. Assisted Excitation of Activations: A Learning Technique to Improve Object Detectors, *Mohammad Mahdi Derakhshani, Saeed Masoudnia, Amir Hossein Shaker, Omid Mersa, Mohammad Amin Sadeghi, Mohammad Rastegari, Babak N. Araabi*
27. Exploiting Edge Features for Graph Neural Networks, *Liyu Gong, Qiang Cheng*
28. Propagation Mechanism for Deep and Wide Neural Networks, *Dejiang Xu, Mong Li Lee, Wynne Hsu*
29. Catastrophic Child's Play: Easy to Perform, Hard to Defend Adversarial Attacks, *Chih-Hui Ho, Brandon Leung, Erik Sandström, Yen Chang, Nuno Vasconcelos*
30. Embedding Complementary Deep Networks for Image Classification, *Qiuyu Chen, Wei Zhang, Jun Yu, Jianping Fan*
40. Shape Robust Text Detection With Progressive Scale Expansion Network, *Wenhai Wang, Enze Xie, Xiang Li, Wenbo Hou, Tong Lu, Gang Yu, Shuai Shao*
41. Dual Encoding for Zero-Example Video Retrieval, *Jianfeng Dong, Xirong Li, Chaoxi Xu, Shouling Ji, Yuan He, Gang Yang, Xun Wang*
42. MaxpoolNMS: Getting Rid of NMS Bottlenecks in Two-Stage Object Detectors, *Lile Cai, Bin Zhao, Zhe Wang, Jie Lin, Chuan Sheng Foo, Mohamed Sabry Aly, Vijay Chandrasekhar*
43. Character Region Awareness for Text Detection, *Youngmin Baek, Bado Lee, Dongyoon Han, Sangdoo Yun, Hwalsuk Lee*
44. Effective Aesthetics Prediction With Multi-Level Spatially Pooled Features, *Vlad Hosu, Bastian Goldlücke, Dietmar Saupe*
45. Attentive Region Embedding Network for Zero-Shot Learning, *Guo-Sen Xie, Li Liu, Xiaobo Jin, Fan Zhu, Zheng Zhang, Jie Qin, Yazhou Yao, Ling Shao*
46. Explicit Spatial Encoding for Deep Local Descriptors, *Arun Mukundan, Giorgos Tolias, Ondřej Chum*
47. Panoptic Segmentation, *Alexander Kirillov, Kaiming He, Ross Girshick, Carsten Rother, Piotr Dollár*
48. You Reap What You Sow: Using Videos to Generate High Precision Object Proposals for Weakly-Supervised Object Detection, *Krishna Kumar Singh, Yong Jae Lee*
49. Explore-Exploit Graph Traversal for Image Retrieval, *Cheng Chang, Guangwei Yu, Chundi Liu, Maksims Volkovs*
50. Dissimilarity Coefficient Based Weakly Supervised Object Detection, *Aditya Arun, C.V. Jawahar, M. Pawan Kumar*
51. Kernel Transformer Networks for Compact Spherical Convolution, *Yu-Chuan Su, Kristen Grauman*
52. Object Detection With Location-Aware Deformable Convolution and Backward Attention Filtering, *Chen Zhang, Joohee Kim*
53. Variational Prototyping-Encoder: One-Shot Learning With Prototypical Images, *Junsik Kim, Tae-Hyun Oh, Seokju Lee, Fei Pan, In So Kweon*
54. Unsupervised Domain Adaptation Using Feature-Whitening and Consensus Loss, *Subhankar Roy, Aliaksandr Siarohin, Enver Sangineto, Samuel Rota Bulò, Nicu Sebe, Elisa Ricci*

Recognition

31. Deep Multimodal Clustering for Unsupervised Audiovisual Learning, *Di Hu, Feiping Nie, Xuelong Li*
32. Dense Classification and Implanting for Few-Shot Learning, *Yann Lifchitz, Yannis Avrithis, Sylvaine Picard, Andrei Bursuc*
33. Class-Balanced Loss Based on Effective Number of Samples, *Yin Cui, Menglin Jia, Tsung-Yi Lin, Yang Song, Serge Belongie*
34. Discovering Visual Patterns in Art Collections With Spatially-Consistent Feature Learning, *Xi Shen, Alexei A. Efros, Mathieu Aubry*
35. Min-Max Statistical Alignment for Transfer Learning, *Samitha Herath, Mehrtaash Harandi, Basura Fernando, Richard Nock*
36. Spatial-Aware Graph Relation Network for Large-Scale Object Detection, *Hang Xu, Chenhan Jiang, Xiaodan Liang, Zhenguo Li*
37. Deformable ConvNets V2: More Deformable, Better Results, *Xizhou Zhu, Han Hu, Stephen Lin, Jifeng Dai*
38. Interaction-And-Aggregation Network for Person Re-Identification, *Ruibing Hou, Bingpeng Ma, Hong Chang, Xinqian Gu, Shiguang Shan, Xilin Chen*
39. Rare Event Detection Using Disentangled Representation Learning, *Ryuhei Hamaguchi, Ken Sakurada, Ryosuke Nakamura*

Segmentation, Grouping, & Shape

55. UPSNet: A Unified Panoptic Segmentation Network, *Yuwen Xiong, Renjie Liao, Hengshuang Zhao, Rui Hu, Min Bai, Ersin Yumer, Raquel Urtasun*
56. JSIS3D: Joint Semantic-Instance Segmentation of 3D Point Clouds With Multi-Task Pointwise Networks and Multi-Value Conditional Random Fields, *Quang-Hieu Pham, Thanh Nguyen, Binh-Son Hua, Gemma Roig, Sai-Kit Yeung*
57. Instance Segmentation by Jointly Optimizing Spatial Embeddings and Clustering Bandwidth, *Davy Neven, Bert De Brabandere, Marc Proesmans, Luc Van Gool*
58. DeepCO3: Deep Instance Co-Segmentation by Co-Peak Search and Co-Saliency Detection, *Kuang-Jui Hsu, Yen-Yu Lin, Yung-Yu Chuang*
59. Improving Semantic Segmentation via Video Propagation and Label Relaxation, *Yi Zhu, Karan Sapra, Fitsum A. Reda, Kevin J. Shih, Shawn Newsam, Andrew Tao, Bryan Catanzaro*
60. Accel: A Corrective Fusion Network for Efficient Semantic Segmentation on Video, *Samvit Jain, Xin Wang, Joseph E. Gonzalez*

61. Shape2Motion: Joint Analysis of Motion Parts and Attributes From 3D Shapes, *Xiaogang Wang, Bin Zhou, Yahao Shi, Xiaowu Chen, Qinqing Zhao, Kai Xu*
 62. Semantic Correlation Promoted Shape-Variant Context for Segmentation, *Henghui Ding, Xudong Jiang, Bing Shuai, Ai Qun Liu, Gang Wang*
 63. Relation-Shape Convolutional Neural Network for Point Cloud Analysis, *Yongcheng Liu, Bin Fan, Shiming Xiang, Chunhong Pan*
 64. Enhancing Diversity of Defocus Blur Detectors via Cross-Ensemble Network, *Wenda Zhao, Bowen Zheng, Qihua Lin, Huchuan Lu*
 65. BubbleNets: Learning to Select the Guidance Frame in Video Object Segmentation by Deep Sorting Frames, *Brent A. Griffin, Jason J. Corso*
 66. Collaborative Global-Local Networks for Memory-Efficient Segmentation of Ultra-High Resolution Images, *Wuyang Chen, Ziyu Jiang, Zhangyang Wang, Kexin Cui, Xiaoning Qian*
 67. Efficient Parameter-Free Clustering Using First Neighbor Relations, *Saqib Sarfraz, Vivek Sharma, Rainer Stiefelhagen*
 68. Learning Personalized Modular Network Guided by Structured Knowledge, *Xiaodan Liang*
 69. A Generative Appearance Model for End-To-End Video Object Segmentation, *Joakim Johnander, Martin Danelljan, Emil Brissman, Fahad Shahbaz Khan, Michael Felsberg*
 70. FEELVOS: Fast End-To-End Embedding Learning for Video Object Segmentation, *Paul Voigtlaender, Yuning Chai, Florian Schroff, Hartwig Adam, Bastian Leibe, Liang-Chieh Chen*
 71. PartNet: A Recursive Part Decomposition Network for Fine-Grained and Hierarchical Shape Segmentation, *Fenggen Yu, Kun Liu, Yan Zhang, Chenyang Zhu, Kai Xu*
 72. Learning Multi-Class Segmentations From Single-Class Datasets, *Konstantin Dmitriev, Arie E. Kaufman*
 73. Convolutional Recurrent Network for Road Boundary Extraction, *Justin Liang, Namdar Homayounfar, Wei-Chiu Ma, Shenlong Wang, Raquel Urtasun*
 74. DFANet: Deep Feature Aggregation for Real-Time Semantic Segmentation, *Hanchao Li, Pengfei Xiong, Haoqiang Fan, Jian Sun*
 75. A Cross-Season Correspondence Dataset for Robust Semantic Segmentation, *Måns Larsson, Erik Stenborg, Lars Hammarstrand, Marc Pollefeys, Torsten Sattler, Fredrik Kahl*
 76. ManTra-Net: Manipulation Tracing Network for Detection and Localization of Image Forgeries With Anomalous Features, *Yue Wu, Wael AbdAlmageed, Premkumar Natarajan*
- Statistics, Physics, Theory, & Datasets**
77. Divergence Triangle for Joint Training of Generator Model, Energy-Based Model, and Inferential Model, *Tian Han, Erik Nijkamp, Xiaolin Fang, Mitch Hill, Song-Chun Zhu, Ying Nian Wu*
 78. Image Deformation Meta-Networks for One-Shot Learning, *Zitian Chen, Yanwei Fu, Yu-Xiong Wang, Lin Ma, Wei Liu, Martial Hebert*
 79. Online High Rank Matrix Completion, *Jicong Fan, Madeleine Udell*
 80. Multispectral Imaging for Fine-Grained Recognition of Powders on Complex Backgrounds, *Tiancheng Zhi, Bernardo R. Pires, Martial Hebert, Srinivasa G. Narasimhan*
 81. ContactDB: Analyzing and Predicting Grasp Contact via Thermal Imaging, *Samarth Brahmhatt, Cusuh Ham, Charles C. Kemp, James Hays*
 82. Robust Subspace Clustering With Independent and Piecewise Identically Distributed Noise Modeling, *Yuanman Li, Jiantao Zhou, Xianwei Zheng, Jinyu Tian, Yuan Yan Tang*
 83. What Correspondences Reveal About Unknown Camera and Motion Models? *Thomas Probst, Ajad Chhatkuli, Danda Pani Paudel, Luc Van Gool*
 84. Self-Calibrating Deep Photometric Stereo Networks, *Guanying Chen, Kai Han, Boxin Shi, Yasuyuki Matsushita, Kwan-Yee K. Wong*
 85. Argoverse: 3D Tracking and Forecasting With Rich Maps, *Ming-Fang Chang, John Lambert, Patsorn Sangkloy, Jagjeet Singh, Slawomir Bak, Andrew Hartnett, De Wang, Peter Carr, Simon Lucey, Deva Ramanan, James Hays*
 86. Side Window Filtering, *Hui Yin, Yuanhao Gong, Guoping Qiu*
 87. Defense Against Adversarial Images Using Web-Scale Nearest-Neighbor Search, *Abhimanyu Dubey, Laurens van der Maaten, Zeki Yalniz, Yixuan Li, Dhruv Mahajan*
 88. Incremental Object Learning From Contiguous Views, *Stefan Stojanov, Samarth Mishra, Ngoc Anh Thai, Nikhil Dhanda, Ahmad Humayun, Chen Yu, Linda B. Smith, James M. Rehg*
 89. IP102: A Large-Scale Benchmark Dataset for Insect Pest Recognition, *Xiaoping Wu, Chi Zhan, Yu-Kun Lai, Ming-Ming Cheng, Jufeng Yang*
 90. CityFlow: A City-Scale Benchmark for Multi-Target Multi-Camera Vehicle Tracking and Re-Identification, *Zheng Tang, Milind Naphade, Ming-Yu Liu, Xiaodong Yang, Stan Birchfield, Shuo Wang, Ratnesh Kumar, David Anastasiu, Jenq-Neng Hwang*
 91. Social-IQ: A Question Answering Benchmark for Artificial Social Intelligence, *Amir Zadeh, Michael Chan, Paul Pu Liang, Edmund Tong, Louis-Philippe Morency*
 92. On Zero-Shot Recognition of Generic Objects, *Tristan Hascoet, Yasuo Aiki, Tetsuya Takiguchi*
 93. Explicit Bias Discovery in Visual Question Answering Models, *Varun Manjunatha, Nirat Saini, Larry S. Davis*
 94. REPAIR: Removing Representation Bias by Dataset Resampling, *Yi Li, Nuno Vasconcelos*
 95. Label Efficient Semi-Supervised Learning via Graph Filtering, *Qimai Li, Xiao-Ming Wu, Han Liu, Xiaotong Zhang, Zhichao Guan*
 96. MVTec AD — A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection, *Paul Bergmann, Michael Fauser, David Sattlegger, Carsten Steger*
 97. ABC: A Big CAD Model Dataset for Geometric Deep Learning, *Sebastian Koch, Albert Matveev, Zhongshi Jiang, Francis Williams, Alexey Artemov, Evgeny Burnaev, Marc Alexa, Denis Zorin, Daniele Panozzo*
 98. Tightness-Aware Evaluation Protocol for Scene Text Detection, *Yuliang Liu, Lianwen Jin, Zecheng Xie, Canjie Luo, Shuaitao Zhang, Lele Xie*
- 3D Multiview**
99. PointConv: Deep Convolutional Networks on 3D Point Clouds, *Wenxuan Wu, Zhongang Qi, Li Fuxin*
 100. Octree Guided CNN With Spherical Kernels for 3D Point Clouds, *Huan Lei, Naveed Akhtar, Ajmal Mian*
 101. VITAMIN-E: Visual Tracking and Mapping With Extremely Dense Feature Points, *Masashi Yokozuka, Shuji Oishi, Simon Thompson, Atsuhiko Banno*
 102. Conditional Single-View Shape Generation for Multi-View Stereo Reconstruction, *Yi Wei, Shaohui Liu, Wang Zhao, Jiwen Lu*

103. Learning to Adapt for Stereo, *Alessio Tonioni, Oscar Rahnama, Thomas Joy, Luigi Di Stefano, Thalaisyasingam Ajanthan, Philip H.S. Torr*
104. 3D Appearance Super-Resolution With Deep Learning, *Yawei Li, Vagia Tsiminaki, Radu Timofte, Marc Pollefeys, Luc Van Gool*
105. Radial Distortion Triangulation, *Zuzana Kukelova, Viktor Larsson*
106. Robust Point Cloud Based Reconstruction of Large-Scale Outdoor Scenes, *Ziquan Lan, Zi Jian Yew, Gim Hee Lee*

3D Single View & RGBD

107. Minimal Solvers for Mini-Loop Closures in 3D Multi-Scan Alignment, *Pedro Miraldo, Surojit Saha, Srikumar Ramalingam*
108. Volumetric Capture of Humans With a Single RGBD Camera via Semi-Parametric Learning, *Rohit Pandey, Anastasia Tkach, Shuoran Yang, Pavel Pidlypenskyi, Jonathan Taylor, Ricardo Martin-Brualla, Andrea Tagliasacchi, George Papandreou, Philip Davidson, Cem Keskin, Shahram Izadi, Sean Fanello*
109. Joint Face Detection and Facial Motion Retargeting for Multiple Faces, *Bindita Chaudhuri, Noranart Vesdapunt, Baoyuan Wang*
110. Monocular Depth Estimation Using Relative Depth Maps, *Jae-Han Lee, Chang-Su Kim*
111. Unsupervised Primitive Discovery for Improved 3D Generative Modeling, *Salman H. Khan, Yulan Guo, Munawar Hayat, Nick Barnes*
112. Learning to Explore Intrinsic Saliency for Stereoscopic Video, *Qiudan Zhang, Xu Wang, Shiqi Wang, Shikai Li, Sam Kwong, Jianmin Jiang*
113. Spherical Regression: Learning Viewpoints, Surface Normals and 3D Rotations on N-Spheres, *Shuai Liao, Efstratios Gavves, Cees G. M. Snoek*
114. Refine and Distill: Exploiting Cycle-Inconsistency and Knowledge Distillation for Unsupervised Monocular Depth Estimation, *Andrea Pilzer, Stéphane Lathuilière, Nicu Sebe, Elisa Ricci*
115. Learning View Priors for Single-View 3D Reconstruction, *Hiroharu Kato, Tatsuya Harada*
116. Geometry-Aware Symmetric Domain Adaptation for Monocular Depth Estimation, *Shanshan Zhao, Huan Fu, Mingming Gong, Dacheng Tao*
117. Learning Monocular Depth Estimation Infusing Traditional Stereo Knowledge, *Fabio Tosi, Filippo Aleotti, Matteo Poggi, Stefano Mattoccia*
118. SIGNet: Semantic Instance Aided Unsupervised 3D Geometry Perception, *Yue Meng, Yongxi Lu, Aman Raj, Samuel Sunarjo, Rui Guo, Tara Javidi, Gaurav Bansal, Dinesh Bharadia*

Face & Body

119. 3D Guided Fine-Grained Face Manipulation, *Zhenglin Geng, Chen Cao, Sergey Tulyakov*
120. Neuro-Inspired Eye Tracking With Eye Movement Dynamics, *Kang Wang, Hui Su, Qiang Ji*
121. Facial Emotion Distribution Learning by Exploiting Low-Rank Label Correlations Locally, *Xiuyi Jia, Xiang Zheng, Weiwei Li, Changqing Zhang, Zechao Li*
122. Unsupervised Face Normalization With Extreme Pose and Expression in the Wild, *Yichen Qian, Weihong Deng, Jiani Hu*
123. Semantic Component Decomposition for Face Attribute Manipulation, *Ying-Cong Chen, Xiaohui Shen, Zhe Lin, Xin Lu, I-Ming Pao, Jiaya Jia*
124. R³ Adversarial Network for Cross Model Face Recognition, *Ken Chen, Yichao Wu, Haoyu Qin, Ding Liang, Xuebo Liu, Junjie Yan*

125. Disentangling Latent Hands for Image Synthesis and Pose Estimation, *Linlin Yang, Angela Yao*
126. Generating Multiple Hypotheses for 3D Human Pose Estimation With Mixture Density Network, *Chen Li, Gim Hee Lee*
127. CrossInfoNet: Multi-Task Information Sharing Based Hand Pose Estimation, *Kuo Du, Xiangbo Lin, Yi Sun, Xiaohong Ma*
128. P2SGrad: Refined Gradients for Optimizing Deep Face Models, *Xiao Zhang, Rui Zhao, Junjie Yan, Mengya Gao, Yu Qiao, Xiaogang Wang, Hongsheng Li*

Action & Video

129. Action Recognition From Single Timestamp Supervision in Untrimmed Videos, *Davide Moltisanti, Sanja Fidler, Dima Damen*
130. Time-Conditioned Action Anticipation in One Shot, *QiuHong Ke, Mario Fritz, Bernt Schiele*
131. Dance With Flow: Two-In-One Stream Action Detection, *Jiaojiao Zhao, Cees G. M. Snoek*
132. Representation Flow for Action Recognition, *AJ Piergiovanni, Michael S. Ryoo*
133. LSTA: Long Short-Term Attention for Egocentric Action Recognition, *Swathikiran Sudhakaran, Sergio Escalera, Oswald Lanz*
134. Learning Actor Relation Graphs for Group Activity Recognition, *Jianchao Wu, Limin Wang, Li Wang, Jie Guo, Gangshan Wu*
135. A Structured Model for Action Detection, *Yubo Zhang, Pavel Tokmakov, Martial Hebert, Cordelia Schmid*
136. Out-Of-Distribution Detection for Generalized Zero-Shot Action Recognition, *Devraj Mandal, Sanath Narayan, Sai Kumar Dwivedi, Vikram Gupta, Shuaib Ahmed, Fahad Shahbaz Khan, Ling Shao*

Motion & Biometrics

137. Object Discovery in Videos as Foreground Motion Clustering, *Christopher Xie, Yu Xiang, Zaid Harchaoui, Dieter Fox*
138. Towards Natural and Accurate Future Motion Prediction of Humans and Animals, *Zhenguang Liu, Shuang Wu, Shuyuan Jin, Qi Liu, Shijian Lu, Roger Zimmermann, Li Cheng*
139. Automatic Face Aging in Videos via Deep Reinforcement Learning, *Chi Nhan Duong, Khoa Luu, Kha Gia Quach, Nghia Nguyen, Eric Patterson, Tien D. Bui, Ngan Le*
140. Multi-Adversarial Discriminative Deep Domain Generalization for Face Presentation Attack Detection, *Rui Shao, Xiangyuan Lan, Jiawei Li, Pong C. Yuen*

Synthesis

141. A Content Transformation Block for Image Style Transfer, *Dmytro Kotovenko, Artsiom Sanakoyeu, Pingchuan Ma, Sabine Lang, Björn Ommer*
142. BeautyGlow: On-Demand Makeup Transfer Framework With Reversible Generative Network, *Hung-Jen Chen, Ka-Ming Hui, Szu-Yu Wang, Li-Wu Tsao, Hong-Han Shuai, Wen-Huang Cheng*
143. Style Transfer by Relaxed Optimal Transport and Self-Similarity, *Nicholas Kolkin, Jason Salavon, Gregory Shakhnarovich*
144. Inserting Videos Into Videos, *Donghoon Lee, Tomas Pfister, Ming-Hsuan Yang*
145. Learning Image and Video Compression Through Spatial-Temporal Energy Compaction, *Zhengxue Cheng, Heming Sun, Masaru Takeuchi, Jiro Katto*

146. Event-Based High Dynamic Range Image and Very High Frame Rate Video Generation Using Conditional Generative Adversarial Networks, *Lin Wang, S. Mohammad Mostafavi I., Yo-Sung Ho, Kuk-Jin Yoon*
147. Enhancing TripleGAN for Semi-Supervised Conditional Instance Synthesis and Classification, *Si Wu, Guangchang Deng, Jichang Li, Rui Li, Zhiwen Yu, Hau-San Wong*

Computational Photography & Graphics

148. Capture, Learning, and Synthesis of 3D Speaking Styles, *Daniel Cudeiro, Timo Bolkart, Cassidy Laidlaw, Anurag Ranjan, Michael J. Black*
149. Nesti-Net: Normal Estimation for Unstructured 3D Point Clouds Using Convolutional Neural Networks, *Yizhak Ben-Shabat, Michael Lindenbaum, Anath Fischer*
150. Ray-Space Projection Model for Light Field Camera, *Qi Zhang, Jinbo Ling, Qing Wang, Jingyi Yu*
151. Deep Geometric Prior for Surface Reconstruction, *Francis Williams, Teseo Schneider, Claudio Silva, Denis Zorin, Joan Bruna, Daniele Panozzo*
152. Analysis of Feature Visibility in Non-Line-Of-Sight Measurements, *Xiaochun Liu, Sebastian Bauer, Andreas Velten*
153. Hyperspectral Imaging With Random Printed Mask, *Yuanyuan Zhao, Hui Guo, Zhan Ma, Xun Cao, Tao Yue, Xuemei Hu*
154. All-Weather Deep Outdoor Lighting Estimation, *Jinsong Zhang, Kalyan Sunkavalli, Yannick Hold-Geoffroy, Sunil Hadap, Jonathan Eisenman, Jean-François Lalonde*

Low-Level & Optimization

155. A Variational EM Framework With Adaptive Edge Selection for Blind Motion Deblurring, *Liuge Yang, Hui Ji*
156. Viewport Proposal CNN for 360° Video Quality Assessment, *Chen Li, Mai Xu, Lai Jiang, Shanyi Zhang, Xiaoming Tao*
157. Beyond Gradient Descent for Regularized Segmentation Losses, *Dmitrii Marin, Meng Tang, Ismail Ben Ayed, Yuri Boykov*
158. MAGSAC: Marginalizing Sample Consensus, *Daniel Barath, Jiří Matas, Jana Noskova*
159. Understanding and Visualizing Deep Visual Saliency Models, *Sen He, Hamed R. Tavakoli, Ali Borji, Yang Mi, Nicolas Pugeault*
160. Divergence Prior and Vessel-Tree Reconstruction, *Zhongwen Zhang, Dmitrii Marin, Egor Chesakov, Marc Moreno Maza, Maria Drangova, Yuri Boykov*
161. Unsupervised Domain-Specific Deblurring via Disentangled Representations, *Boyu Lu, Jun-Cheng Chen, Rama Chellappa*
162. Douglas-Rachford Networks: Learning Both the Image Prior and Data Fidelity Terms for Blind Image Deconvolution, *Raied Aljadaany, Dipan K. Pal, Marios Savvides*
163. Speed Invariant Time Surface for Learning to Detect Corner Points With Event-Based Cameras, *Jacques Manderscheid, Amos Sironi, Nicolas Bourdis, Davide Migliore, Vincent Lepetit*
164. Training Deep Learning Based Image Denoisers From Undersampled Measurements Without Ground Truth and Without Image Prior, *Maguiyi Zhussip, Shakarim Soltanayev, Se Young Chun*
165. A Variational Pan-Sharpener With Local Gradient Constraints, *Xueyang Fu, Zihuang Lin, Yue Huang, Xinghao Ding*

Scenes & Representation

166. F-VAEGAN-D2: A Feature Generating Framework for Any-Shot Learning, *Yongqin Xian, Saurabh Sharma, Bernt Schiele, Zeynep Akata*

167. Sliced Wasserstein Discrepancy for Unsupervised Domain Adaptation, *Chen-Yu Lee, Tanmay Batra, Mohammad Haris Baig, Daniel Ulbricht*
168. Graph Attention Convolution for Point Cloud Semantic Segmentation, *Lei Wang, Yuchun Huang, Yaolin Hou, Shenman Zhang, Jie Shan*
169. Normalized Diversification, *Shaohui Liu, Xiao Zhang, Jianqiao Wangni, Jianbo Shi*
170. Learning to Localize Through Compressed Binary Maps, *Xinkai Wei, Ioan Andrei Bârsan, Shenlong Wang, Julieta Martinez, Raquel Urtasun*
171. A Parametric Top-View Representation of Complex Road Scenes, *Ziyan Wang, Buyu Liu, Samuel Schuster, Manmohan Chandraker*
172. Self-Supervised Spatiotemporal Learning via Video Clip Order Prediction, *Dejing Xu, Jun Xiao, Zhou Zhao, Jian Shao, Di Xie, Yueting Zhuang*
173. Superquadrics Revisited: Learning 3D Shape Parsing Beyond Cuboids, *Despoina Paschalidou, Ali Osman Ulusoy, Andreas Geiger*
174. Unsupervised Disentangling of Appearance and Geometry by Deformable Generator Network, *Xianglei Xing, Tian Han, Ruiqi Gao, Song-Chun Zhu, Ying Nian Wu*
175. Self-Supervised Representation Learning by Rotation Feature Decoupling, *Zeyu Feng, Chang Xu, Dacheng Tao*
176. Weakly Supervised Deep Image Hashing Through Tag Embeddings, *Vijetha Gattupalli, Yaoxin Zhuo, Baoxin Li*
177. Improved Road Connectivity by Joint Learning of Orientation and Segmentation, *Anil Batra, Suriya Singh, Guan Pang, Saikat Basu, C.V. Jawahar, Manohar Paluri*
178. Deep Supervised Cross-Modal Retrieval, *Liangli Zhen, Peng Hu, Xu Wang, Dezhong Peng*
179. A Theoretically Sound Upper Bound on the Triplet Loss for Improving the Efficiency of Deep Distance Metric Learning, *Thanh-Toan Do, Toan Tran, Ian Reid, Vijay Kumar, Tuan Hoang, Gustavo Carneiro*
180. Data Representation and Learning With Graph Diffusion-Embedding Networks, *Bo Jiang, Doudou Lin, Jin Tang, Bin Luo*

Language & Reasoning

181. Video Relationship Reasoning Using Gated Spatio-Temporal Energy Graph, *Yao-Hung Hubert Tsai, Santosh Divvala, Louis-Philippe Morency, Ruslan Salakhutdinov, Ali Farhadi*
182. Image-Question-Answer Synergistic Network for Visual Dialog, *Dalu Guo, Chang Xu, Dacheng Tao*
183. Not All Frames Are Equal: Weakly-Supervised Video Grounding With Contextual Similarity and Visual Clustering Losses, *Jing Shi, Jia Xu, Boqing Gong, Chenliang Xu*
184. Inverse Cooking: Recipe Generation From Food Images, *Amaia Salvador, Michal Drozdal, Xavier Giro-i-Nieto, Adriana Romero*
185. Adversarial Semantic Alignment for Improved Image Captions, *Pierre Dognin, Igor Melnyk, Youssef Mroueh, Jerret Ross, Tom Sercu*
186. Answer Them All! Toward Universal Visual Question Answering Models, *Robik Shrestha, Kushal Kafle, Christopher Kanan*
187. Unsupervised Multi-Modal Neural Machine Translation, *Yuanhang Su, Kai Fan, Nguyen Bach, C.-C. Jay Kuo, Fei Huang*
188. Multi-Task Learning of Hierarchical Vision-Language Representation, *Duy-Kien Nguyen, Takayuki Okatani*

1320–1520 Setup for Poster Session 3-2P (Exhibit Hall)**1330–1520 Oral Session 3-2A: Deep Learning**
(Terrace Theater)

Papers in this session are in Poster Session 3-2P (Posters 1–18)

Chairs: Judy Hoffman (*Facebook AI Research; Georgia Tech*)
Philipp Kraehenbuehl (*Univ. of Texas at Austin*)**Format** (5 min. presentation; 3 min. group questions/3 papers)

1. [1330] Practical Full Resolution Learned Lossless Image Compression, *Fabian Mentzer, Eirikur Agustsson, Michael Tschannen, Radu Timofte, Luc Van Gool*
2. [1335] Image-To-Image Translation via Group-Wise Deep Whitening-And-Coloring Transformation, *Wonwoong Cho, Sungha Choi, David Keetae Park, Inkyu Shin, Jaegul Choo*
3. [1340] Max-Sliced Wasserstein Distance and Its Use for GANs, *Ishan Deshpande, Yuan-Ting Hu, Ruoyu Sun, Ayis Pyrros, Nasir Siddiqui, Sanmi Koyejo, Zhizhen Zhao, David Forsyth, Alexander G. Schwing*
4. [1348] Meta-Learning With Differentiable Convex Optimization, *Kwonjoon Lee, Subhransu Maji, Avinash Ravichandran, Stefano Soatto*
5. [1353] RePr: Improved Training of Convolutional Filters, *Aaditya Prakash, James Storer, Dinei Florencio, Cha Zhang*
6. [1358] Tangent-Normal Adversarial Regularization for Semi-Supervised Learning, *Bing Yu, Jingfeng Wu, Jinwen Ma, Zhanxing Zhu*
7. [1406] Auto-Encoding Scene Graphs for Image Captioning, *Xu Yang, Kaihua Tang, Hanwang Zhang, Jianfei Cai*
8. [1411] Fast, Diverse and Accurate Image Captioning Guided by Part-Of-Speech, *Aditya Deshpande, Jyoti Aneja, Liwei Wang, Alexander G. Schwing, David Forsyth*
9. [1416] Attention Branch Network: Learning of Attention Mechanism for Visual Explanation, *Hiroshi Fukui, Tsubasa Hirakawa, Takayoshi Yamashita, Hironobu Fujiyoshi*
10. [1424] Cascaded Projection: End-To-End Network Compression and Acceleration, *Breton Minnehan, Andreas Savakis*
11. [1429] DeepCaps: Going Deeper With Capsule Networks, *Jathushan Rajasegaran, Vinoj Jayasundara, Sandaru Jayasekara, Hirunima Jayasekara, Suranga Seneviratne, Ranga Rodrigo*
12. [1434] FBNet: Hardware-Aware Efficient ConvNet Design via Differentiable Neural Architecture Search, *Bichen Wu, Xiaoliang Dai, Peizhao Zhang, Yanghan Wang, Fei Sun, Yiming Wu, Yuandong Tian, Peter Vajda, Yangqing Jia, Kurt Keutzer*
13. [1442] APDrawingGAN: Generating Artistic Portrait Drawings From Face Photos With Hierarchical GANs, *Ran Yi, Yong-Jin Liu, Yu-Kun Lai, Paul L. Rosin*
14. [1447] Constrained Generative Adversarial Networks for Interactive Image Generation, *Eric Heim*
15. [1452] WarpGAN: Automatic Caricature Generation, *Yichun Shi, Debayan Deb, Anil K. Jain*
16. [1500] Explainability Methods for Graph Convolutional Neural Networks, *Phillip E. Pope, Soheil Kolouri, Mohammad Rostami, Charles E. Martin, Heiko Hoffmann*
17. [1505] A Generative Adversarial Density Estimator, *M. Ehsan Abbasnejad, Qinfeng Shi, Anton van den Hengel, Lingqiao Liu*

18. [1510] SoDeep: A Sorting Deep Net to Learn Ranking Loss Surrogates, *Martin Engilberge, Louis Chevallier, Patrick Pérez, Matthieu Cord*

1330–1520 Oral Session 3-2B: Face & Body
(Grand Ballroom)

Papers in this session are in Poster Session 3-2P (Posters 92–109)

Chairs: Simon Lucey (*Carnegie Mellon Univ.*)
Dimitris Samaras (*Stony Brook Univ.*)**Format** (5 min. presentation; 3 min. group questions/3 papers)

1. [1330] High-Quality Face Capture Using Anatomical Muscles, *Michael Bao, Matthew Cong, Stéphane Grabli, Ronald Fedkiw*
2. [1335] FML: Face Model Learning From Videos, *Ayush Tewari, Florian Bernard, Pablo Garrido, Gaurav Bharaj, Mohamed Elgharib, Hans-Peter Seidel, Patrick Pérez, Michael Zollhöfer, Christian Theobalt*
3. [1340] AdaCos: Adaptively Scaling Cosine Logits for Effectively Learning Deep Face Representations, *Xiao Zhang, Rui Zhao, Yu Qiao, Xiaogang Wang, Hongsheng Li*
4. [1348] 3D Hand Shape and Pose Estimation From a Single RGB Image, *Liuhaog Ge, Zhou Ren, Yuncheng Li, Zehao Xue, Yingying Wang, Jianfei Cai, Junsong Yuan*
5. [1353] 3D Hand Shape and Pose From Images in the Wild, *Adnane Boukhayma, Rodrigo de Bem, Philip H.S. Torr*
6. [1358] Self-Supervised 3D Hand Pose Estimation Through Training by Fitting, *Chengde Wan, Thomas Probst, Luc Van Gool, Angela Yao*
7. [1406] CrowdPose: Efficient Crowded Scenes Pose Estimation and a New Benchmark, *Jiefeng Li, Can Wang, Hao Zhu, Yihuan Mao, Hao-Shu Fang, Cewu Lu*
8. [1411] Towards Social Artificial Intelligence: Nonverbal Social Signal Prediction in a Triadic Interaction, *Hanbyul Joo, Tomas Simon, Mina Cikara, Yaser Sheikh*
9. [1416] HoloPose: Holistic 3D Human Reconstruction In-The-Wild, *Riza Alp Güler, Iasonas Kokkinos*
10. [1424] Weakly-Supervised Discovery of Geometry-Aware Representation for 3D Human Pose Estimation, *Xipeng Chen, Kwan-Yee Lin, Wentao Liu, Chen Qian, Liang Lin*
11. [1429] In the Wild Human Pose Estimation Using Explicit 2D Features and Intermediate 3D Representations, *Ikhsanul Habibie, Weipeng Xu, Dushyant Mehta, Gerard Pons-Moll, Christian Theobalt*
12. [1434] Slim DensePose: Thrifty Learning From Sparse Annotations and Motion Cues, *Natalia Neverova, James Thewlis, Riza Alp Güler, Iasonas Kokkinos, Andrea Vedaldi*
13. [1442] Self-Supervised Representation Learning From Videos for Facial Action Unit Detection, *Yong Li, Jiabei Zeng, Shiguang Shan, Xilin Chen*
14. [1447] Combining 3D Morphable Models: A Large Scale Face-And-Head Model, *Stylios Ploumpis, Haoyang Wang, Nick Pears, William A. P. Smith, Stefanos Zafeiriou*
15. [1452] Boosting Local Shape Matching for Dense 3D Face Correspondence, *Zhenfeng Fan, Xiyuan Hu, Chen Chen, Silong Peng*
16. [1500] Unsupervised Part-Based Disentangling of Object Shape and Appearance, *Dominik Lorenz, Leonard Bereska, Timo Milbich, Björn Ommer*

17. [1505] Monocular Total Capture: Posing Face, Body, and Hands in the Wild, *Donglai Xiang, Hanbyul Joo, Yaser Sheikh*
18. [1510] Expressive Body Capture: 3D Hands, Face, and Body From a Single Image, *Georgios Pavlakos, Vasileios Choutas, Nima Ghorbani, Timo Bolkart, Ahmed A. Osman, Dimitrios Tzionas, Michael J. Black*

1330–1520 Oral Session 3-2C: Low-Level & Optimization (Promenade Ballroom)

Papers in this session are in Poster Session 3-2P (Posters 147–164)

Chairs: Sing Bing Kang (*Zillow Group*)
Ce Liu (*Google*)

Format (5 min. presentation; 3 min. group questions/3 papers)

1. [1330] Neural RGB→D Sensing: Depth and Uncertainty From a Video Camera, *Chao Liu, Jinwei Gu, Kihwan Kim, Srinivasa G. Narasimhan, Jan Kautz*
2. [1335] DAVANet: Stereo Deblurring With View Aggregation, *Shangchen Zhou, Jiawei Zhang, Wangmeng Zuo, Haozhe Xie, Jinshan Pan, Jimmy S. Ren*
3. [1340] DVC: An End-To-End Deep Video Compression Framework, *Guo Lu, Wanli Ouyang, Dong Xu, Xiaoyun Zhang, Chunlei Cai, Zhiyong Gao*
4. [1348] SOSNet: Second Order Similarity Regularization for Local Descriptor Learning, *Yurun Tian, Xin Yu, Bin Fan, Fuchao Wu, Huub Heijnen, Vasileios Balntas*
5. [1353] “Double-DIP”: Unsupervised Image Decomposition via Coupled Deep-Image-Priors, *Yosef Gandelsman, Assaf Shocher, Michal Irani*
6. [1358] Unprocessing Images for Learned Raw Denoising, *Tim Brooks, Ben Mildenhall, Tianfan Xue, Jiawen Chen, Dillon Sharlet, Jonathan T. Barron*
7. [1406] Residual Networks for Light Field Image Super-Resolution, *Shuo Zhang, Youfang Lin, Hao Sheng*
8. [1411] Modulating Image Restoration With Continual Levels via Adaptive Feature Modification Layers, *Jingwen He, Chao Dong, Yu Qiao*
9. [1416] Second-Order Attention Network for Single Image Super-Resolution, *Tao Dai, Jianrui Cai, Yongbing Zhang, Shu-Tao Xia, Lei Zhang*
10. [1424] Devil Is in the Edges: Learning Semantic Boundaries From Noisy Annotations, *David Acuna, Amlan Kar, Sanja Fidler*
11. [1429] Path-Invariant Map Networks, *Zaiwei Zhang, Zhenxiao Liang, Lemeng Wu, Xiaowei Zhou, Qixing Huang*
12. [1434] FilterReg: Robust and Efficient Probabilistic Point-Set Registration Using Gaussian Filter and Twist Parameterization, *Wei Gao, Russ Tedrake*
13. [1442] Probabilistic Permutation Synchronization Using the Riemannian Structure of the Birkhoff Polytope, *Tolga Birdal, Umut Şimşekli*
14. [1447] Lifting Vectorial Variational Problems: A Natural Formulation Based on Geometric Measure Theory and Discrete Exterior Calculus, *Thomas Möllenhoff, Daniel Cremers*
15. [1452] A Sufficient Condition for Convergences of Adam and RMSProp, *Fangyu Zou, Li Shen, Zequn Jie, Weizhong Zhang, Wei Liu*

16. [1500] Guaranteed Matrix Completion Under Multiple Linear Transformations, *Chao Li, Wei He, Longhao Yuan, Zhun Sun, Qibin Zhao*
17. [1505] MAP Inference via Block-Coordinate Frank-Wolfe Algorithm, *Paul Swoboda, Vladimir Kolmogorov*
18. [1510] A Convex Relaxation for Multi-Graph Matching, *Paul Swoboda, Dagmar Kainmüller, Ashkan Mokarian, Christian Theobalt, Florian Bernard*

1520–1620 Afternoon Break (Exhibit Hall)

1520–1800 Demos (Exhibit Hall)

- Deep ChArUco: Dark ChArUco Marker Pose Estimation, *Danying Hu, Daniel DeTone, Tomasz Malisiewicz (Magic Leap)*
- Real-Time Monocular Depth Estimation Without GPU, *Matteo Poggi, Fabio Tosi, Filippo Aleotti, Stefano Mattoccia (Univ. of Bologna)*
- A Camera That CNNs: Convolutional Neural Network on a Pixel Processor Array, *Laurie Bose, Piotr Dudek, Walterio Mayol (Univ. of Manchester and Univ. of Bristol)*
- Active Illumination to Increase Visibility Range in Bad Weather Conditions, *Manvi Malik (IBM)*

1520–1800 Exhibits (Exhibit Hall)

- See Exhibits map for list of exhibitors.

1520–1800 Poster Session 3-2P (Exhibit Hall)

Deep Learning

1. Practical Full Resolution Learned Lossless Image Compression, *Fabian Mentzer, Eirikur Agustsson, Michael Tschannen, Radu Timofte, Luc Van Gool*
2. Image-To-Image Translation via Group-Wise Deep Whitening-And-Coloring Transformation, *Wonwoong Cho, Sungha Choi, David Keetae Park, Inkyu Shin, Jaegul Choo*
3. Max-Sliced Wasserstein Distance and Its Use for GANs, *Ishan Deshpande, Yuan-Ting Hu, Ruoyu Sun, Ayis Pyrrhos, Nasir Siddiqui, Sanmi Koyejo, Zhizhen Zhao, David Forsyth, Alexander G. Schwing*
4. Meta-Learning With Differentiable Convex Optimization, *Kwonjoon Lee, Subhransu Maji, Avinash Ravichandran, Stefano Soatto*
5. RePr: Improved Training of Convolutional Filters, *Aaditya Prakash, James Storer, Dinei Florencio, Cha Zhang*
6. Tangent-Normal Adversarial Regularization for Semi-Supervised Learning, *Bing Yu, Jingfeng Wu, Jinwen Ma, Zhanxing Zhu*
7. Auto-Encoding Scene Graphs for Image Captioning, *Xu Yang, Kaihua Tang, Hanwang Zhang, Jianfei Cai*
8. Fast, Diverse and Accurate Image Captioning Guided by Part-Of-Speech, *Aditya Deshpande, Jyoti Aneja, Liwei Wang, Alexander G. Schwing, David Forsyth*
9. Attention Branch Network: Learning of Attention Mechanism for Visual Explanation, *Hiroshi Fukui, Tsubasa Hirakawa, Takayoshi Yamashita, Hironobu Fujiyoshi*
10. Cascaded Projection: End-To-End Network Compression and Acceleration, *Breton Minnehan, Andreas Savakis*

11. DeepCaps: Going Deeper With Capsule Networks, *Jathushan Rajasegaran, Vinoj Jayasundara, Sandaru Jayasekara, Hirunima Jayasekara, Suranga Seneviratne, Ranga Rodrigo*
 12. FBNet: Hardware-Aware Efficient ConvNet Design via Differentiable Neural Architecture Search, *Bichen Wu, Xiaoliang Dai, Peizhao Zhang, Yanghan Wang, Fei Sun, Yiming Wu, Yuandong Tian, Peter Vajda, Yangqing Jia, Kurt Keutzer*
 13. APDrawingGAN: Generating Artistic Portrait Drawings From Face Photos With Hierarchical GANs, *Ran Yi, Yong-Jin Liu, Yu-Kun Lai, Paul L. Rosin*
 14. Constrained Generative Adversarial Networks for Interactive Image Generation, *Eric Heim*
 15. WarpGAN: Automatic Caricature Generation, *Yichun Shi, Debayan Deb, Anil K. Jain*
 16. Explainability Methods for Graph Convolutional Neural Networks, *Phillip E. Pope, Soheil Kolouri, Mohammad Rostami, Charles E. Martin, Heiko Hoffmann*
 17. A Generative Adversarial Density Estimator, *M. Ehsan Abbasnejad, Qinfeng Shi, Anton van den Hengel, Lingqiao Liu*
 18. SoDeep: A Sorting Deep Net to Learn Ranking Loss Surrogates, *Martin Engilberge, Louis Chevallier, Patrick Pérez, Matthieu Cord*
 19. Pixel-Adaptive Convolutional Neural Networks, *Hang Su, Varun Jampani, Deqing Sun, Orazio Gallo, Erik Learned-Miller, Jan Kautz*
 20. Single-Frame Regularization for Temporally Stable CNNs, *Gabriel Eilertsen, Rafal K. Mantiuk, Jonas Unger*
 21. An End-To-End Network for Generating Social Relationship Graphs, *Arushi Goel, Keng Teck Ma, Cheston Tan*
 22. Meta-Learning Convolutional Neural Architectures for Multi-Target Concrete Defect Classification With the CONcrete DEfect BRidge IMage Dataset, *Martin Mundt, Sagnik Majumder, Sreenivas Murali, Panagiotis Panetsos, Visvanathan Ramesh*
 23. ECC: Platform-Independent Energy-Constrained Deep Neural Network Compression via a Bilinear Regression Model, *Haichuan Yang, Yuhao Zhu, Ji Liu*
 24. SeerNet: Predicting Convolutional Neural Network Feature-Map Sparsity Through Low-Bit Quantization, *Shijie Cao, Lingxiao Ma, Wencong Xiao, Chen Zhang, Yunxin Liu, Lintao Zhang, Lanshun Nie, Zhi Yang*
 25. Defending Against Adversarial Attacks by Randomized Diversification, *Olga Taran, Shideh Rezaeifar, Taras Holotyak, Slava Voloshynovskiy*
 26. Rob-GAN: Generator, Discriminator, and Adversarial Attacker, *Xuanqing Liu, Cho-Jui Hsieh*
 27. Learning From Noisy Labels by Regularized Estimation of Annotator Confusion, *Ryutaro Tanno, Ardavan Saeedi, Swami Sankaranarayanan, Daniel C. Alexander, Nathan Silberman*
 28. Task-Free Continual Learning, *Rahaf Aljundi, Klaas Kelchtermans, Tinne Tuytelaars*
 29. Importance Estimation for Neural Network Pruning, *Pavlo Molchanov, Arun Mallya, Stephen Tyree, Iuri Frosio, Jan Kautz*
 30. Detecting Overfitting of Deep Generative Networks via Latent Recovery, *Ryan Webster, Julien Rabin, Loïc Simon, Frédéric Jurie*
 31. Coloring With Limited Data: Few-Shot Colorization via Memory Augmented Networks, *Seungjoo Yoo, Hyojin Bahng, Sunghyo Chung, Junsoo Lee, Jaehyuk Chang, Jaegul Choo*
 32. Characterizing and Avoiding Negative Transfer, *Zirui Wang, Zihang Dai, Barnabás Póczos, Jaime Carbonell*
 33. Building Efficient Deep Neural Networks With Unitary Group Convolutions, *Ritchie Zhao, Yuwei Hu, Jordan Dotzel, Christopher De Sa, Zhiru Zhang*
 34. Semi-Supervised Learning With Graph Learning-Convolutional Networks, *Bo Jiang, Ziyang Zhang, Doudou Lin, Jin Tang, Bin Luo*
 35. Learning to Remember: A Synaptic Plasticity Driven Framework for Continual Learning, *Oleksiy Ostapenko, Mihai Puscas, Tassilo Klein, Patrick Jähnechen, Moin Nabi*
 36. AIRD: Adversarial Learning Framework for Image Repurposing Detection, *Ayush Jaiswal, Yue Wu, Wael AbdAlmageed, Iacopo Masi, Premkumar Natarajan*
 37. A Kernelized Manifold Mapping to Diminish the Effect of Adversarial Perturbations, *Saeid Asgari Taghanaki, Kumar Abhishek, Shekoofeh Azizi, Ghassan Hamarneh*
 38. Trust Region Based Adversarial Attack on Neural Networks, *Zhewei Yao, Amir Gholami, Peng Xu, Kurt Keutzer, Michael W. Mahoney*
 39. PEPSI : Fast Image Inpainting With Parallel Decoding Network, *Min-cheol Sagong, Yong-goo Shin, Seung-wook Kim, Seung Park, Sung-jea Ko*
 40. Model-Blind Video Denoising via Frame-To-Frame Training, *Thibaud Ehret, Axel Davy, Jean-Michel Morel, Gabriele Facciolo, Pablo Arias*
 41. End-To-End Efficient Representation Learning via Cascading Combinatorial Optimization, *Yeonwoo Jeong, Yoonsung Kim, Hyun Oh Song*
 42. Sim-Real Joint Reinforcement Transfer for 3D Indoor Navigation, *Fengda Zhu, Linchao Zhu, Yi Yang*
 43. ChamNet: Towards Efficient Network Design Through Platform-Aware Model Adaptation, *Xiaoliang Dai, Peizhao Zhang, Bichen Wu, Hongxu Yin, Fei Sun, Yanghan Wang, Marat Dukhan, Yunqing Hu, Yiming Wu, Yangqing Jia, Peter Vajda, Matt Uyttendaele, Niraj K. Jha*
 44. Regularizing Activation Distribution for Training Binarized Deep Networks, *Ruizhou Ding, Ting-Wu Chin, Zeye Liu, Diana Marculescu*
 45. Robustness Verification of Classification Deep Neural Networks via Linear Programming, *Wang Lin, Zhengfeng Yang, Xin Chen, Qingye Zhao, Xiangkun Li, Zhiming Liu, Jifeng He*
 46. Additive Adversarial Learning for Unbiased Authentication, *Jian Liang, Yuren Cao, Chenbin Zhang, Shiyu Chang, Kun Bai, Zenglin Xu*
 47. Simultaneously Optimizing Weight and Quantizer of Ternary Neural Network Using Truncated Gaussian Approximation, *Zhezhi He, Deliang Fan*
 48. Adversarial Defense by Stratified Convolutional Sparse Coding, *Bo Sun, Nian-Hsuan Tsai, Fangchen Liu, Ronald Yu, Hao Su*
- Recognition**
49. Exploring Object Relation in Mean Teacher for Cross-Domain Detection, *Qi Cai, Yingwei Pan, Chong-Wah Ngo, Xinmei Tian, Lingyu Duan, Ting Yao*
 50. Hierarchical Disentanglement of Discriminative Latent Features for Zero-Shot Learning, *Bin Tong, Chao Wang, Martin Klinkigt, Yoshiyuki Kobayashi, Yuuichi Nonaka*
 51. R²GAN: Cross-Modal Recipe Retrieval With Generative Adversarial Network, *Bin Zhu, Chong-Wah Ngo, Jingjing Chen, Yanbin Hao*

52. Rethinking Knowledge Graph Propagation for Zero-Shot Learning, *Michael Kampffmeyer, Yinbo Chen, Xiaodan Liang, Hao Wang, Yuja Zhang, Eric P. Xing*
53. Learning to Learn Image Classifiers With Visual Analogy, *Linjun Zhou, Peng Cui, Shiqiang Yang, Wenwu Zhu, Qi Tian*
54. Where's Wally Now? Deep Generative and Discriminative Embeddings for Novelty Detection, *Philippe Burlina, Neil Joshi, I-Jeng Wang*
55. Weakly Supervised Image Classification Through Noise Regularization, *Mengying Hu, Hu Han, Shiguang Shan, Xilin Chen*
56. Data-Driven Neuron Allocation for Scale Aggregation Networks, *Yi Li, Zhanghui Kuang, Yimin Chen, Wayne Zhang*
57. Graphical Contrastive Losses for Scene Graph Parsing, *Ji Zhang, Kevin J. Shih, Ahmed Elgammal, Andrew Tao, Bryan Catanzaro*
58. Deep Transfer Learning for Multiple Class Novelty Detection, *Pramuditha Perera, Vishal M. Patel*
59. QATM: Quality-Aware Template Matching for Deep Learning, *Jiaxin Cheng, Yue Wu, Wael AbdAlmageed, Premkumar Natarajan*
60. Retrieval-Augmented Convolutional Neural Networks Against Adversarial Examples, *Jake Zhao (Junbo), Kyunghyun Cho*
61. Learning Cross-Modal Embeddings With Adversarial Networks for Cooking Recipes and Food Images, *Hao Wang, Doyen Sahoo, Chenghao Liu, Ee-peng Lim, Steven C. H. Hoi*
62. FastDraw: Addressing the Long Tail of Lane Detection by Adapting a Sequential Prediction Network, *Jonah Philion*
63. Weakly Supervised Video Moment Retrieval From Text Queries, *Niluthpol Chowdhury Mithun, Sujoy Paul, Amit K. Roy-Chowdhury*

Segmentation, Grouping, & Shape

64. Content-Aware Multi-Level Guidance for Interactive Instance Segmentation, *Soumajit Majumder, Angela Yao*
65. Greedy Structure Learning of Hierarchical Compositional Models, *Adam Kortylewski, Aleksander Wiczkorek, Mario Wieser, Clemens Blumer, Sonali Parbhoo, Andreas Morel-Forster, Volker Roth, Thomas Vetter*
66. Interactive Full Image Segmentation by Considering All Regions Jointly, *Eirikur Agustsson, Jasper R. R. Uijlings, Vittorio Ferrari*
67. Learning Active Contour Models for Medical Image Segmentation, *Xu Chen, Bryan M. Williams, Srinivasa R. Vallabhaneni, Gabriela Czanner, Rachel Williams, Yalin Zheng*
68. Customizable Architecture Search for Semantic Segmentation, *Yiheng Zhang, Zhaofan Qiu, Jingen Liu, Ting Yao, Dong Liu, Tao Mei*

Statistics, Physics, Theory, & Datasets

69. Local Features and Visual Words Emerge in Activations, *Oriane Siméoni, Yannis Avrithis, Ondřej Chum*
70. Hyperspectral Image Super-Resolution With Optimized RGB Guidance, *Ying Fu, Tao Zhang, Yinqiang Zheng, Debing Zhang, Hua Huang*
71. Adaptive Confidence Smoothing for Generalized Zero-Shot Learning, *Yuval Atzmon, Gal Chechik*
72. PMS-Net: Robust Haze Removal Based on Patch Map for Single Images, *Wei-Ting Chen, Jian-Jiun Ding, Sy-Yen Kuo*
73. Deep Spherical Quantization for Image Search, *Sepehr Eghbali, Ladan Tahvildari*
74. Large-Scale Interactive Object Segmentation With Human Annotators, *Rodrigo Benenson, Stefan Popov, Vittorio Ferrari*

75. A Poisson-Gaussian Denoising Dataset With Real Fluorescence Microscopy Images, *Yide Zhang, Yin hao Zhu, Evan Nichols, Qingfei Wang, Siyuan Zhang, Cody Smith, Scott Howard*
76. Task Agnostic Meta-Learning for Few-Shot Learning, *Muhammad Abdullah Jamal, Guo-Jun Qi*
77. Progressive Ensemble Networks for Zero-Shot Recognition, *Meng Ye, Yuhong Guo*
78. Direct Object Recognition Without Line-Of-Sight Using Optical Coherence, *Xin Lei, Liangyu He, Yixuan Tan, Ken Xingze Wang, Xinggang Wang, Yihan Du, Shanhuai Fan, Zongfu Yu*
79. Atlas of Digital Pathology: A Generalized Hierarchical Histological Tissue Type-Annotated Database for Deep Learning, *Mahdi S. Hosseini, Lyndon Chan, Gabriel Tse, Michael Tang, Jun Deng, Sajad Norouzi, Corwyn Rowsell, Konstantinos N. Plataniotis, Savvas Damaskinos*

3D Multiview

80. Perturbation Analysis of the 8-Point Algorithm: A Case Study for Wide FoV Cameras, *Thiago L. T. da Silveira, Claudio R. Jung*
81. Robustness of 3D Deep Learning in an Adversarial Setting, *Matthew Wicker, Marta Kwiatkowska*
82. SceneCode: Monocular Dense Semantic Reconstruction Using Learned Encoded Scene Representations, *Shuaifeng Zhi, Michael Bloesch, Stefan Leutenegger, Andrew J. Davison*
83. StereoDRNet: Dilated Residual StereoNet, *Rohan Chabra, Julian Straub, Christopher Sweeney, Richard Newcombe, Henry Fuchs*
84. The Alignment of the Spheres: Globally-Optimal Spherical Mixture Alignment for Camera Pose Estimation, *Dylan Campbell, Lars Petersson, Laurent Kneip, Hongdong Li, Stephen Gould*

3D Single View & RGBD

85. Learning Joint Reconstruction of Hands and Manipulated Objects, *Yana Hasson, Gül Varol, Dimitrios Tzionas, Igor Kalevatykh, Michael J. Black, Ivan Laptev, Cordelia Schmid*
86. Deep Single Image Camera Calibration With Radial Distortion, *Manuel López, Roger Marí, Pau Gargallo, Yubin Kuang, Javier Gonzalez-Jimenez, Gloria Haro*
87. CAM-Convs: Camera-Aware Multi-Scale Convolutions for Single-View Depth, *Jose M. Facil, Benjamin Ummenhofer, Huizhong Zhou, Luis Montesano, Thomas Brox, Javier Civera*
88. Translate-to-Recognize Networks for RGB-D Scene Recognition, *Dapeng Du, Limin Wang, Huiling Wang, Kai Zhao, Gangshan Wu*
89. Re-Identification Supervised Texture Generation, *Jian Wang, Yunshan Zhong, Yachun Li, Chi Zhang, Yichen Wei*
90. Action4D: Online Action Recognition in the Crowd and Clutter, *Quanzeng You, Hao Jiang*
91. Monocular 3D Object Detection Leveraging Accurate Proposals and Shape Reconstruction, *Jason Ku, Alex D. Pon, Steven L. Waslander*

Face & Body

92. High-Quality Face Capture Using Anatomical Muscles, *Michael Bao, Matthew Cong, Stéphane Grabli, Ronald Fedkiw*
93. FML: Face Model Learning From Videos, *Ayush Tewari, Florian Bernard, Pablo Garrido, Gaurav Bharaj, Mohamed Elgharib, Hans-Peter Seidel, Patrick Pérez, Michael Zollhöfer, Christian Theobalt*

94. AdaCos: Adaptively Scaling Cosine Logits for Effectively Learning Deep Face Representations, *Xiao Zhang, Rui Zhao, Yu Qiao, Xiaogang Wang, Hongsheng Li*
 95. 3D Hand Shape and Pose Estimation From a Single RGB Image, *Liuhaog Ge, Zhou Ren, Yuncheng Li, Zehao Xue, Yingying Wang, Jianfei Cai, Junsong Yuan*
 96. 3D Hand Shape and Pose From Images in the Wild, *Adnane Boukhayma, Rodrigo de Bem, Philip H.S. Torr*
 97. Self-Supervised 3D Hand Pose Estimation Through Training by Fitting, *Chengde Wan, Thomas Probst, Luc Van Gool, Angela Yao*
 98. CrowdPose: Efficient Crowded Scenes Pose Estimation and a New Benchmark, *Jiefeng Li, Can Wang, Hao Zhu, Yihuan Mao, Hao-Shu Fang, Cewu Lu*
 99. Towards Social Artificial Intelligence: Nonverbal Social Signal Prediction in a Triadic Interaction, *Hanbyul Joo, Tomas Simon, Mina Cikara, Yaser Sheikh*
 100. HoloPose: Holistic 3D Human Reconstruction In-The-Wild, *Riza Alp Güler, Iasonas Kokkinos*
 101. Weakly-Supervised Discovery of Geometry-Aware Representation for 3D Human Pose Estimation, *Xipeng Chen, Kwan-Yee Lin, Wentao Liu, Chen Qian, Liang Lin*
 102. In the Wild Human Pose Estimation Using Explicit 2D Features and Intermediate 3D Representations, *Ikhsanul Habibie, Weipeng Xu, Dushyant Mehta, Gerard Pons-Moll, Christian Theobalt*
 103. Slim DensePose: Thrifty Learning From Sparse Annotations and Motion Cues, *Natalia Neverova, James Thewlis, Riza Alp Güler, Iasonas Kokkinos, Andrea Vedaldi*
 104. Self-Supervised Representation Learning From Videos for Facial Action Unit Detection, *Yong Li, Jiabei Zeng, Shiguang Shan, Xilin Chen*
 105. Combining 3D Morphable Models: A Large Scale Face-And-Head Model, *Stylios Ploumpis, Haoyang Wang, Nick Pears, William A. P. Smith, Stefanos Zafeiriou*
 106. Boosting Local Shape Matching for Dense 3D Face Correspondence, *Zhenfeng Fan, Xiyuan Hu, Chen Chen, Silong Peng*
 107. Unsupervised Part-Based Disentangling of Object Shape and Appearance, *Dominik Lorenz, Leonard Bereska, Timo Milbich, Björn Ommer*
 108. Monocular Total Capture: Posing Face, Body, and Hands in the Wild, *Donglai Xiang, Hanbyul Joo, Yaser Sheikh*
 109. Expressive Body Capture: 3D Hands, Face, and Body From a Single Image, *Georgios Pavlakos, Vasileios Choutas, Nima Ghorbani, Timo Bolkart, Ahmed A. A. Osman, Dimitrios Tzionas, Michael J. Black*
 110. Attribute-Aware Face Aging With Wavelet-Based Generative Adversarial Networks, *Yunfan Liu, Qi Li, Zhenan Sun*
 111. Noise-Tolerant Paradigm for Training Face Recognition CNNs, *Wei Hu, Yangyu Huang, Fan Zhang, Ruirui Li*
 112. Low-Rank Laplacian-Uniform Mixed Model for Robust Face Recognition, *Jiayu Dong, Huicheng Zheng, Lina Lian*
 113. Generalizing Eye Tracking With Bayesian Adversarial Learning, *Kang Wang, Rui Zhao, Hui Su, Qiang Ji*
 114. Local Relationship Learning With Person-Specific Shape Regularization for Facial Action Unit Detection, *Xuesong Niu, Hu Han, Songfan Yang, Yan Huang, Shiguang Shan*
 115. Point-To-Pose Voting Based Hand Pose Estimation Using Residual Permutation Equivariant Layer, *Shile Li, Dongheui Lee*
 116. Improving Few-Shot User-Specific Gaze Adaptation via Gaze Redirection Synthesis, *Yu Yu, Gang Liu, Jean-Marc Odobez*
 117. AdaptiveFace: Adaptive Margin and Sampling for Face Recognition, *Hao Liu, Xiangyu Zhu, Zhen Lei, Stan Z. Li*
 118. Disentangled Representation Learning for 3D Face Shape, *Zi-Hang Jiang, Qianyi Wu, Keyu Chen, Juyong Zhang*
 119. LBS Autoencoder: Self-Supervised Fitting of Articulated Meshes to Point Clouds, *Chun-Liang Li, Tomas Simon, Jason Saragih, Barnabás Póczos, Yaser Sheikh*
 120. PiPaf: Composite Fields for Human Pose Estimation, *Sven Kreiss, Lorenzo Bertoni, Alexandre Alahi*
- Action & Video**
121. TACNet: Transition-Aware Context Network for Spatio-Temporal Action Detection, *Lin Song, Shiwei Zhang, Gang Yu, Hongbin Sun*
 122. Learning Regularity in Skeleton Trajectories for Anomaly Detection in Videos, *Romero Morais, Vuong Le, Truyen Tran, Budhaditya Saha, Moussa Mansour, Svetha Venkatesh*
 123. Local Temporal Bilinear Pooling for Fine-Grained Action Parsing, *Yan Zhang, Siyu Tang, Krikamol Muandet, Christian Jarvers, Heiko Neumann*
 124. Improving Action Localization by Progressive Cross-Stream Cooperation, *Rui Su, Wanli Ouyang, Luping Zhou, Dong Xu*
 125. Two-Stream Adaptive Graph Convolutional Networks for Skeleton-Based Action Recognition, *Lei Shi, Yifan Zhang, Jian Cheng, Hanqing Lu*
 126. A Neural Network Based on SPD Manifold Learning for Skeleton-Based Hand Gesture Recognition, *Xuan Son Nguyen, Luc Brun, Olivier Lézoray, Sébastien Bougleux*
 127. Large-Scale Weakly-Supervised Pre-Training for Video Action Recognition, *Deepti Ghadiyaram, Du Tran, Dhruv Mahajan*
 128. Learning Spatio-Temporal Representation With Local and Global Diffusion, *Zhaofan Qiu, Ting Yao, Chong-Wah Ngo, Xinmei Tian, Tao Mei*
 129. Unsupervised Learning of Action Classes With Continuous Temporal Embedding, *Anna Kukleva, Hilde Kuehne, Fadime Sener, Jürgen Gall*
 130. Double Nuclear Norm Based Low Rank Representation on Grassmann Manifolds for Clustering, *Xinglin Piao, Yongli Hu, Junbin Gao, Yanfeng Sun, Baocai Yin*
- Motion & Biometrics**
131. SR-LSTM: State Refinement for LSTM Towards Pedestrian Trajectory Prediction, *Pu Zhang, Wanli Ouyang, Pengfei Zhang, Jianru Xue, Nanning Zheng*
 132. Unsupervised Deep Epipolar Flow for Stationary or Dynamic Scenes, *Yiran Zhong, Pan Ji, Jianyuan Wang, Yuchao Dai, Hongdong Li*
 133. An Efficient Schmidt-EKF for 3D Visual-Inertial SLAM, *Patrick Geneva, James Maley, Guoquan Huang*
 134. A Neural Temporal Model for Human Motion Prediction, *Anand Gopalakrishnan, Ankur Mali, Dan Kifer, Lee Giles, Alexander G. Ororbia*
 135. Multi-Agent Tensor Fusion for Contextual Trajectory Prediction, *Tianyang Zhao, Yifei Xu, Mathew Monfort, Wongun Choi, Chris Baker, Yibiao Zhao, Yizhou Wang, Ying Nian Wu*

Synthesis

136. Coordinate-Based Texture Inpainting for Pose-Guided Human Image Generation, *Artur Grigorev, Artem Sevastopolsky, Alexander Vakhitov, Victor Lempitsky*
137. On Stabilizing Generative Adversarial Training With Noise, *Simon Jenni, Paolo Favaro*
138. Self-Supervised GANs via Auxiliary Rotation Loss, *Ting Chen, Xiaohua Zhai, Marvin Ritter, Mario Lucic, Neil Houlsby*
139. Texture Mixer: A Network for Controllable Synthesis and Interpolation of Texture, *Ning Yu, Connelly Barnes, Eli Shechtman, Sohrab Amirghodsi, Michal Lukáč*
140. Object-Driven Text-To-Image Synthesis via Adversarial Training, *Wenbo Li, Pengchuan Zhang, Lei Zhang, Qiuyuan Huang, Xiaodong He, Siwei Lyu, Jianfeng Gao*
141. Zoom-In-To-Check: Boosting Video Interpolation via Instance-Level Discrimination, *Liangzhe Yuan, Yibo Chen, Hantian Liu, Tao Kong, Jianbo Shi*
142. Disentangling Latent Space for VAE by Label Relevant/Irrelevant Dimensions, *Zhilin Zheng, Li Sun*

Computational Photography & Graphics

143. Spectral Reconstruction From Dispersive Blur: A Novel Light Efficient Spectral Imager, *Yuanyuan Zhao, Xuemei Hu, Hui Guo, Zhan Ma, Tao Yue, Xun Cao*
144. Quasi-Unsupervised Color Constancy, *Simone Bianco, Claudio Cusano*
145. Deep Defocus Map Estimation Using Domain Adaptation, *Junyong Lee, Sungkil Lee, Sunghyun Cho, Seungyong Lee*
146. Using Unknown Occluders to Recover Hidden Scenes, *Adam B. Yedidia, Manel Baradad, Christos Thrampoulidis, William T. Freeman, Gregory W. Wornell*

Low-Level & Optimization

147. Neural RGB→D Sensing: Depth and Uncertainty From a Video Camera, *Chao Liu, Jinwei Gu, Kihwan Kim, Srinivasa G. Narasimhan, Jan Kautz*
148. DAVANet: Stereo Deblurring With View Aggregation, *Shangchen Zhou, Jiawei Zhang, Wangmeng Zuo, Haozhe Xie, Jinshan Pan, Jimmy S. Ren*
149. DVC: An End-To-End Deep Video Compression Framework, *Guo Lu, Wanli Ouyang, Dong Xu, Xiaoyun Zhang, Chunlei Cai, Zhiyong Gao*
150. SOSNet: Second Order Similarity Regularization for Local Descriptor Learning, *Yurun Tian, Xin Yu, Bin Fan, Fuchao Wu, Huub Heijnen, Vassileios Balntas*
151. "Double-DIP": Unsupervised Image Decomposition via Coupled Deep-Image-Priors, *Yosef Gandelsman, Assaf Shocher, Michal Irani*
152. Unprocessing Images for Learned Raw Denoising, *Tim Brooks, Ben Mildenhall, Tianfan Xue, Jiawen Chen, Dillon Sharlet, Jonathan T. Barron*
153. Residual Networks for Light Field Image Super-Resolution, *Shuo Zhang, Youfang Lin, Hao Sheng*
154. Modulating Image Restoration With Continual Levels via Adaptive Feature Modification Layers, *Jingwen He, Chao Dong, Yu Qiao*
155. Second-Order Attention Network for Single Image Super-Resolution, *Tao Dai, Jianrui Cai, Yongbing Zhang, Shu-Tao Xia, Lei Zhang*
156. Devil Is in the Edges: Learning Semantic Boundaries From Noisy Annotations, *David Acuna, Amlan Kar, Sanja Fidler*

157. Path-Invariant Map Networks, *Zaiwei Zhang, Zhenxiao Liang, Lemeng Wu, Xiaowei Zhou, Qixing Huang*
158. FilterReg: Robust and Efficient Probabilistic Point-Set Registration Using Gaussian Filter and Twist Parameterization, *Wei Gao, Russ Tedrake*
159. Probabilistic Permutation Synchronization Using the Riemannian Structure of the Birkhoff Polytope, *Tolga Birdal, Umut Şimşekli*
160. Lifting Vectorial Variational Problems: A Natural Formulation Based on Geometric Measure Theory and Discrete Exterior Calculus, *Thomas Möllenhoff, Daniel Cremers*
161. A Sufficient Condition for Convergences of Adam and RMSProp, *Fangyu Zou, Li Shen, Zequn Jie, Weizhong Zhang, Wei Liu*
162. Guaranteed Matrix Completion Under Multiple Linear Transformations, *Chao Li, Wei He, Longhao Yuan, Zhun Sun, Qibin Zhao*
163. MAP Inference via Block-Coordinate Frank-Wolfe Algorithm, *Paul Swoboda, Vladimir Kolmogorov*
164. A Convex Relaxation for Multi-Graph Matching, *Paul Swoboda, Dagmar Kainmüller, Ashkan Mokarian, Christian Theobalt, Florian Bernard*
165. Competitive Collaboration: Joint Unsupervised Learning of Depth, Camera Motion, Optical Flow and Motion Segmentation, *Anurag Ranjan, Varun Jampani, Lukas Balles, Kihwan Kim, Deqing Sun, Jonas Wulff, Michael J. Black*
166. Learning Parallax Attention for Stereo Image Super-Resolution, *Longguang Wang, Yingqian Wang, Zhengfa Liang, Zaiping Lin, Jungang Yang, Wei An, Yulan Guo*
167. Knowing When to Stop: Evaluation and Verification of Conformity to Output-Size Specifications, *Chenglong Wang, Rudy Bunel, Krishnamurthy Dvijotham, Po-Sen Huang, Edward Grefenstette, Pushmeet Kohli*
168. Spatial Attentive Single-Image Deraining With a High Quality Real Rain Dataset, *Tianyu Wang, Xin Yang, Ke Xu, Shaozhe Chen, Qiang Zhang, Rynson W.H. Lau*
169. Focus Is All You Need: Loss Functions for Event-Based Vision, *Guillermo Gallego, Mathias Gehrig, Davide Scaramuzza*
170. Scalable Convolutional Neural Network for Image Compressed Sensing, *Wuzhen Shi, Feng Jiang, Shaohui Liu, Debin Zhao*
171. Event Cameras, Contrast Maximization and Reward Functions: An Analysis, *Timo Stoffregen, Lindsay Kleeman*
172. Convolutional Neural Networks Can Be Deceived by Visual Illusions, *Alexander Gomez-Villa, Adrian Martín, Javier Vazquez-Corral, Marcelo Bertalmio*
173. PDE Acceleration for Active Contours, *Anthony Yezzi, Ganesh Sundaramoorthi, Minas Benyamin*
174. Dichromatic Model Based Temporal Color Constancy for AC Light Sources, *Jun-Sang Yoo, Jong-Ok Kim*
175. Semantic Attribute Matching Networks, *Seungryong Kim, Dongbo Min, Somi Jeong, Sunok Kim, Sangryul Jeon, Kwanghoon Sohn*
176. Skin-Based Identification From Multispectral Image Data Using CNNs, *Takeshi Uemori, Atsushi Ito, Yusuke Moriuchi, Alexander Gatto, Jun Murayama*
177. Large-Scale Distributed Second-Order Optimization Using Kronecker-Factored Approximate Curvature for Deep Convolutional Neural Networks, *Kazuki Osawa, Yohei Tsuji, Yuichiro Ueno, Akira Naruse, Rio Yokota, Satoshi Matsuoka*

Scenes & Representation

178. Putting Humans in a Scene: Learning Affordance in 3D Indoor Environments, *Xueting Li, Sifei Liu, Kihwan Kim, Xiaolong Wang, Ming-Hsuan Yang, Jan Kautz*
179. PIEs: Pose Invariant Embeddings, *Chih-Hui Ho, Pedro Morgado, Amir Persekian, Nuno Vasconcelos*
180. Representation Similarity Analysis for Efficient Task Taxonomy & Transfer Learning, *Kshitij Dwivedi, Gemma Roig*
181. Object Counting and Instance Segmentation With Image-Level Supervision, *Hisham Cholakkal, Guolei Sun, Fahad Shahbaz Khan, Ling Shao*
182. Variational Autoencoders Pursue PCA Directions (by Accident), *Michal Rolínek, Dominik Zietlow, Georg Martius*
183. A Relation-Augmented Fully Convolutional Network for Semantic Segmentation in Aerial Scenes, *Lichao Mou, Yuansheng Hua, Xiao Xiang Zhu*
184. Temporal Transformer Networks: Joint Learning of Invariant and Discriminative Time Warping, *Suhas Lohit, Qiao Wang, Pavan Turaga*
185. PCAN: 3D Attention Map Learning Using Contextual Information for Point Cloud Based Retrieval, *Wenxiao Zhang, Chunxia Xiao*
186. Depth Coefficients for Depth Completion, *Saif Imran, Yunfei Long, Xiaoming Liu, Daniel Morris*
187. Diversify and Match: A Domain Adaptive Representation Learning Paradigm for Object Detection, *Taekyung Kim, Minki Jeong, Seunghyeon Kim, Seokeon Choi, Changick Kim*

Language & Reasoning

188. Good News, Everyone! Context Driven Entity-Aware Captioning for News Images, *Ali Furkan Biten, Lluís Gomez, Marçal Rusiñol, Dimosthenis Karatzas*
189. Multi-Level Multimodal Common Semantic Space for Image-Phrase Grounding, *Hassan Akbari, Svebor Karaman, Surabhi Bhargava, Brian Chen, Carl Vondrick, Shih-Fu Chang*
190. Spatio-Temporal Dynamics and Semantic Attribute Enriched Visual Encoding for Video Captioning, *Nayyer Afaq, Naveed Akhtar, Wei Liu, Syed Zulqarnain Gilani, Ajmal Mian*
191. Pointing Novel Objects in Image Captioning, *Yehao Li, Ting Yao, Yingwei Pan, Hongyang Chao, Tao Mei*
192. Informative Object Annotations: Tell Me Something I Don't Know, *Lior Bracha, Gal Chechik*
193. Engaging Image Captioning via Personality, *Kurt Shuster, Samuel Humeau, Hexiang Hu, Antoine Bordes, Jason Weston*
194. Vision-Based Navigation With Language-Based Assistance via Imitation Learning With Indirect Intervention, *Khanh Nguyen, Debadepta Dey, Chris Brockett, Bill Dolan*
195. TOUCHDOWN: Natural Language Navigation and Spatial Reasoning in Visual Street Environments, *Howard Chen, Alane Suhr, Dipendra Misra, Noah Snaveley, Yoav Artzi*
196. A Simple Baseline for Audio-Visual Scene-Aware Dialog, *Idan Schwartz, Alexander G. Schwing, Tamir Hazan*

Applications, Medical, & Robotics

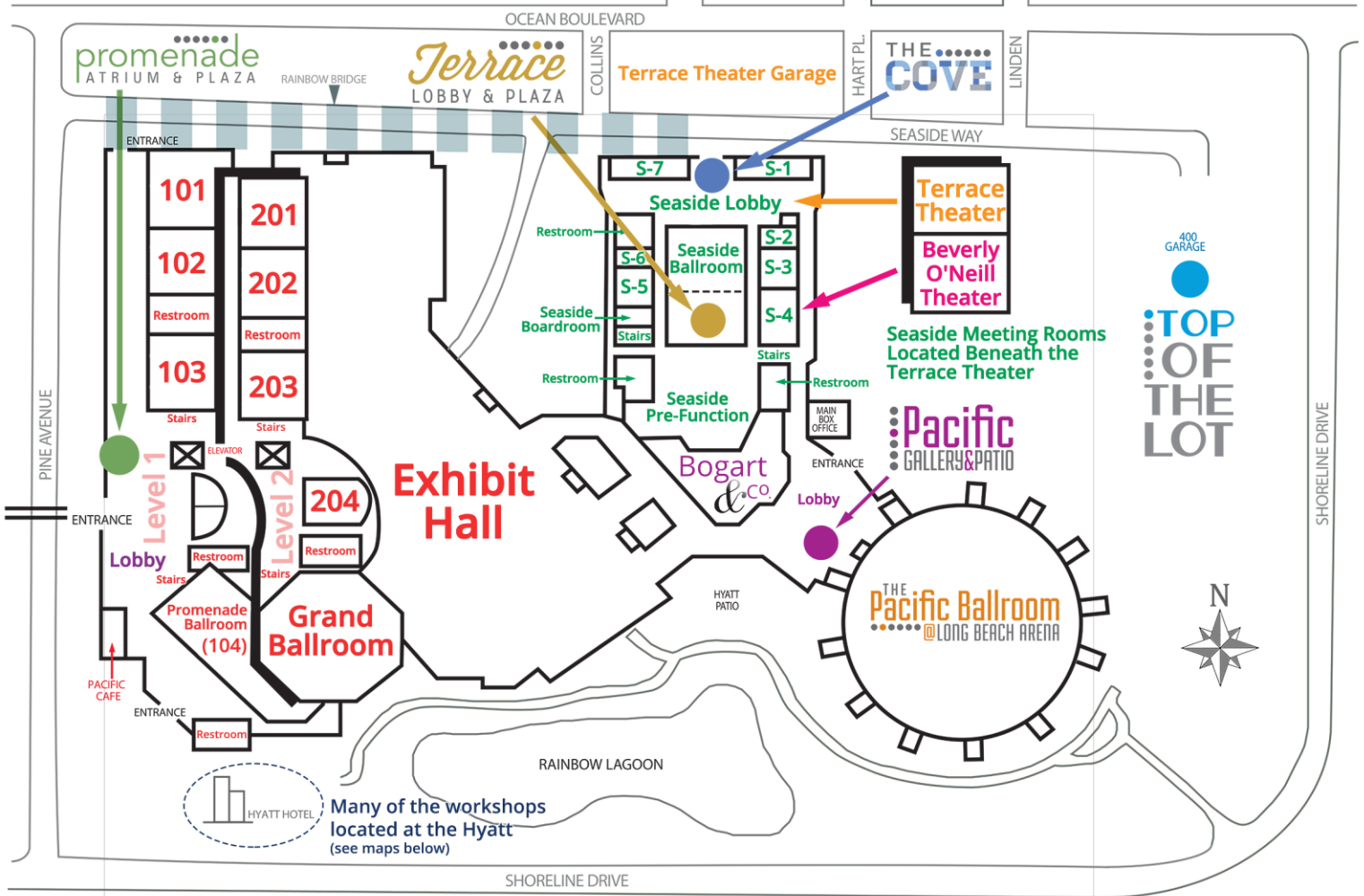
197. End-To-End Learned Random Walker for Seeded Image Segmentation, *Lorenzo Cerrone, Alexander Zeilmann, Fred A. Hamprecht*
198. Efficient Neural Network Compression, *Hyeji Kim, Muhammad Umar Karim Khan, Chong-Min Kyung*

199. Cascaded Generative and Discriminative Learning for Microcalcification Detection in Breast Mammograms, *Fandong Zhang, Ling Luo, Xinwei Sun, Zhen Zhou, Xiuli Li, Yizhou Yu, Yizhou Wang*
200. C3AE: Exploring the Limits of Compact Model for Age Estimation, *Chao Zhang, Shuaicheng Liu, Xun Xu, Ce Zhu*
201. Adaptive Weighting Multi-Field-Of-View CNN for Semantic Segmentation in Pathology, *Hiroki Tokunaga, Yuki Teramoto, Akihiko Yoshizawa, Ryoma Bise*
202. In Defense of Pre-Trained ImageNet Architectures for Real-Time Semantic Segmentation of Road-Driving Images, *Marin Oršić, Ivan Krešo, Petra Bevandić, Siniša Šegvic*
203. Context-Aware Visual Compatibility Prediction, *Guillem Cucurull, Perouz Taslakian, David Vazquez*
204. Sim-To-Real via Sim-To-Sim: Data-Efficient Robotic Grasping via Randomized-To-Canonical Adaptation Networks, *Stephen James, Paul Wohlhart, Mrinal Kalakrishnan, Dmitry Kalashnikov, Alex Irpan, Julian Ibarz, Sergey Levine, Raia Hadsell, Konstantinos Bousmalis*
205. Multiview 2D/3D Rigid Registration via a Point-Of-Interest Network for Tracking and Triangulation, *Haofu Liao, Wei-An Lin, Jiarui Zhang, Jingdan Zhang, Jiebo Luo, S. Kevin Zhou*
206. Context-Aware Spatio-Recurrent Curvilinear Structure Segmentation, *Feigege Wang, Yue Gu, Wenxi Liu, Yuanlong Yu, Shengfeng He, Jia Pan*
207. An Alternative Deep Feature Approach to Line Level Keyword Spotting, *George Retsinas, Georgios Louloudis, Nikolaos Stamatopoulos, Giorgos Sfikas, Basilis Gatos*
208. Dynamics Are Important for the Recognition of Equine Pain in Video, *Sofia Broomé, Karina Bech Gleeerup, Pia Haubro Andersen, Hedvig Kjellström*
209. LaserNet: An Efficient Probabilistic 3D Object Detector for Autonomous Driving, *Gregory P. Meyer, Ankit Laddha, Eric Kee, Carlos Vallespi-Gonzalez, Carl K. Wellington*
210. Machine Vision Guided 3D Medical Image Compression for Efficient Transmission and Accurate Segmentation in the Clouds, *Zihao Liu, Xiaowei Xu, Tao Liu, Qi Liu, Yanzhi Wang, Yiyu Shi, Wujie Wen, Meiping Huang, Haiyun Yuan, Jian Zhuang*
211. PointPillars: Fast Encoders for Object Detection From Point Clouds, *Alex H. Lang, Sourabh Vora, Holger Caesar, Lubing Zhou, Jiong Yang, Oscar Beijbom*
212. Motion Estimation of Non-Holonomic Ground Vehicles From a Single Feature Correspondence Measured Over N Views, *Kun Huang, Yifu Wang, Laurent Kneip*
213. From Coarse to Fine: Robust Hierarchical Localization at Large Scale, *Paul-Edouard Sarlin, Cesar Cadena, Roland Siegwart, Marcin Dymczyk*
214. Large Scale High-Resolution Land Cover Mapping With Multi-Resolution Data, *Caleb Robinson, Le Hou, Kolya Malkin, Rachel Soobitsky, Jacob Czawlytko, Bistra Dilkina, Nebojsa Jojic*
215. Leveraging Heterogeneous Auxiliary Tasks to Assist Crowd Counting, *Muming Zhao, Jian Zhang, Chongyang Zhang, Wenjun Zhang*



LONG BEACH

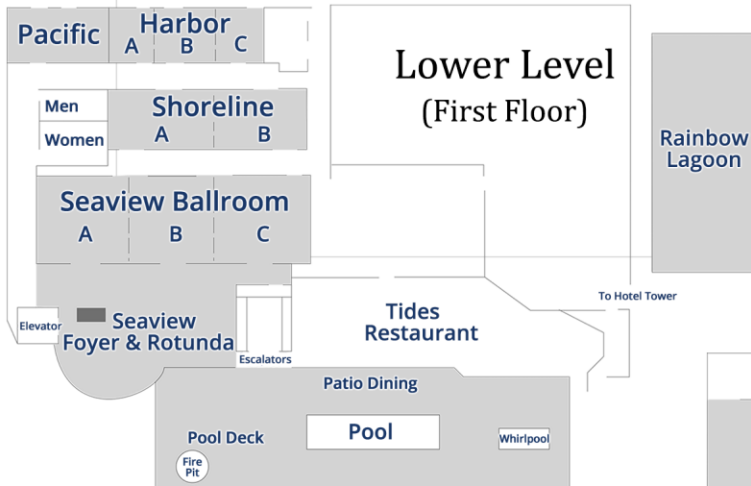
Convention & Entertainment Center



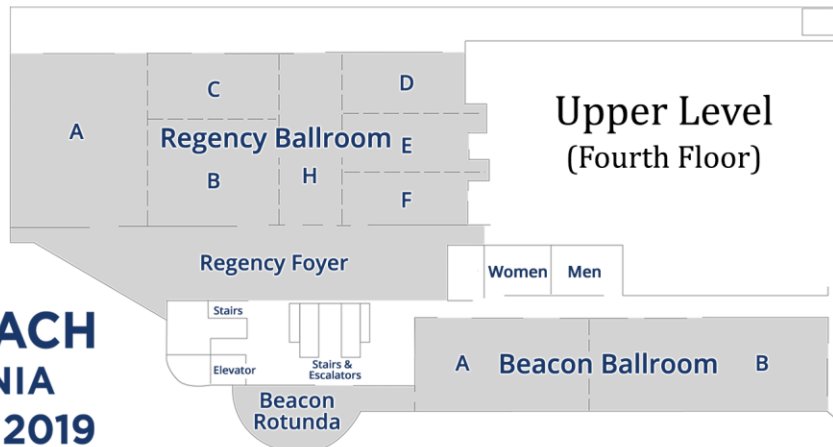
CVPR 2019 Sponsors



IEEE
COMPUTER
SOCIETY



Hyatt Meeting Rooms
(used for many of the workshops)



LONG BEACH CALIFORNIA
June 16-20, 2019

DIAMOND SPONSORS



PLATINUM SPONSORS



GOLD SPONSORS



SILVER SPONSORS

