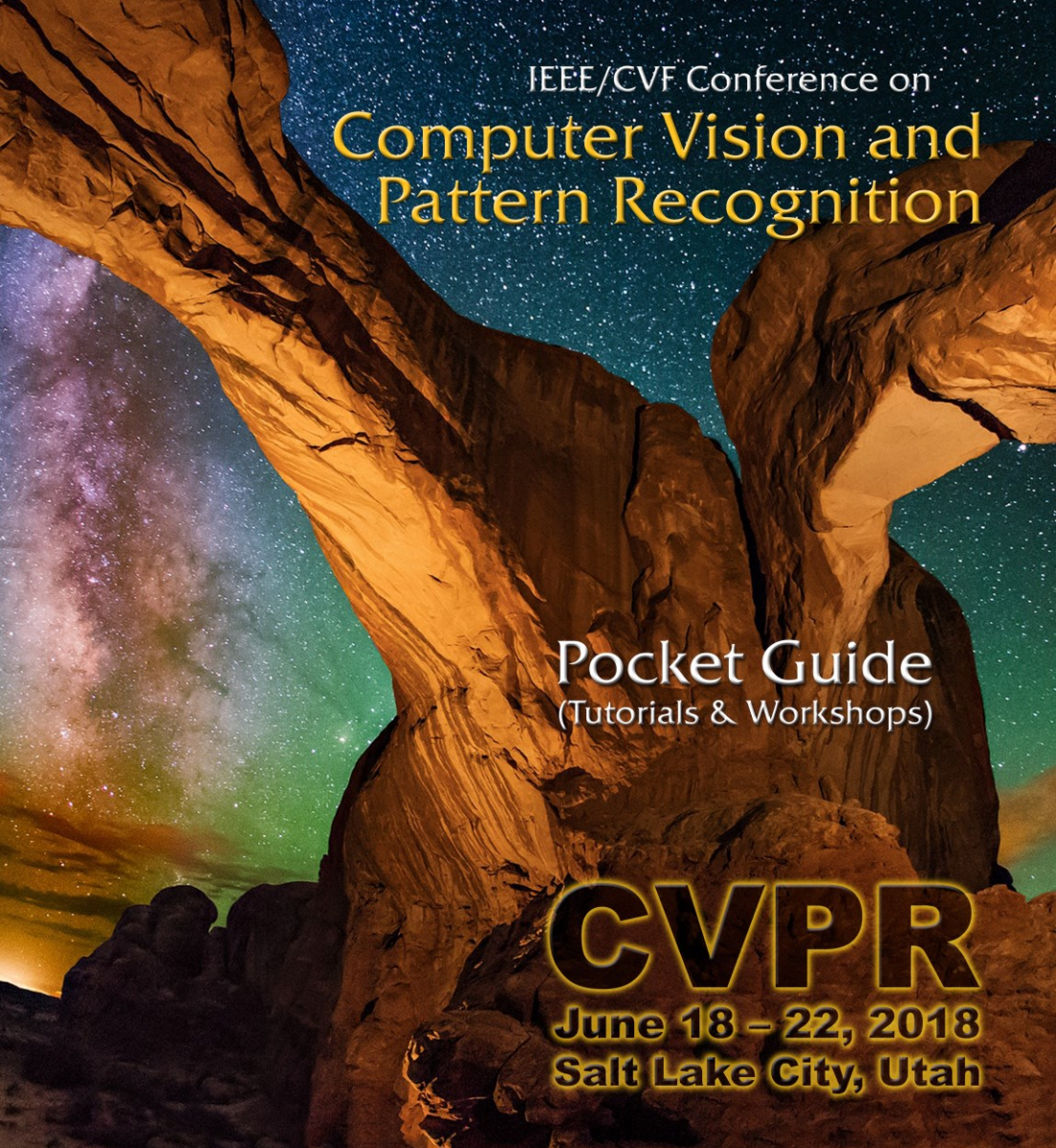
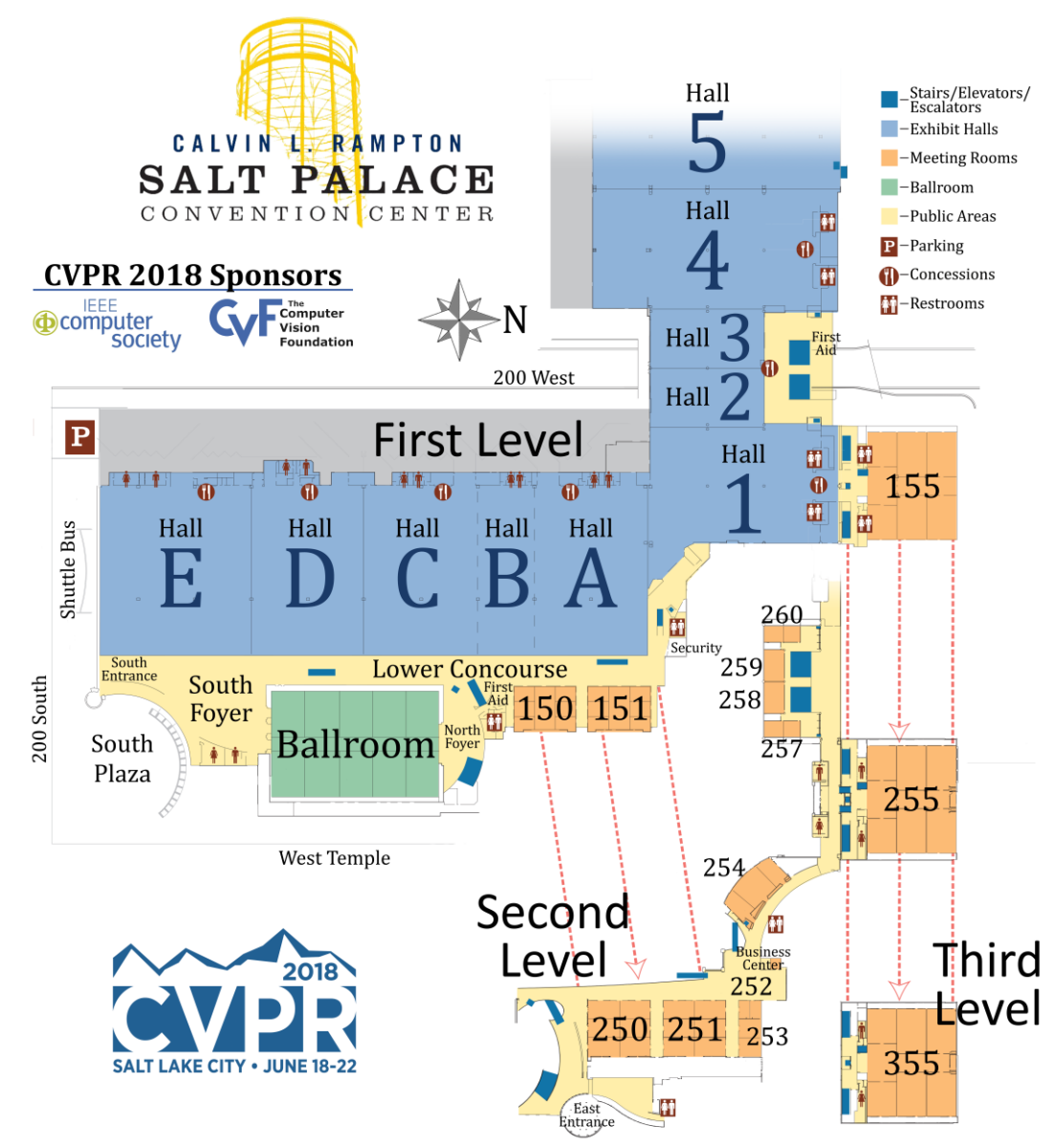

IEEE/CVF Conference on
Computer Vision and Pattern Recognition
Pocket Guide
(Tutorials & Workshops)
CVPR
June 18 – 22, 2018
Salt Lake City, Utah


CALVIN L. RAMPTON
SALT PALACE
CONVENTION CENTER
CVPR 2018 Sponsors
IEEE computer society
CVF The Computer Vision Foundation
N
Stairs/Elevators/ Escalators
Exhibit Halls
Meeting Rooms
Ballroom
Public Areas
Parking
Concessions
Restrooms
Hall 5
Hall 4
Hall 3
Hall 2
Hall 1
First Aid
200 West
First Level
Hall E
Hall D
Hall C
Hall B
Hall A
155
Shuttle Bus
South Entrance
South Foyer
Lower Concourse
Security
260
259
258
257
255
First Aid
150
151
North Foyer
Ballroom
South Plaza
200 South
West Temple
254
Second Level
Business Center
252
250
251
253
East Entrance
Third Level
355
2018 CVPR
SALT LAKE CITY • JUNE 18-22

Message from the General and Program Chairs

Welcome to Salt Lake City and the 31st meeting of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2018). CVPR continues to be a showcase for some of the most exciting advances in computer vision, pattern recognition, machine learning, and artificial intelligence. With a full slate of oral presentations, poster sessions, workshops, tutorials, demos, social functions, and our growing industrial exhibition, this promises to be a week with something for everyone.

As our field continues to grow, so too do our conferences. CVPR 2014 first passed 2,000 attendees, and within three years CVPR 2017 had grown to nearly 5,000, with this year's meeting eclipsing that mark. As attendance has grown, so too has the number of submitted papers. This year's meeting received 3,359 submissions, of which 3,309 were fully reviewed after a handful of administrative rejections. To handle the reviewing process for that many papers, we invited 108 Area Chairs (ACs). This meant including many up-and-coming researchers who had not previously served as ACs and who were paired with more experienced ones. We recruited a pool of 2,385 reviewers, out of which 1,715 were invited to perform reviews. Of these, 575 reviewed the maximum number (10) of assigned papers. One hundred and sixty emergency reviewers graciously provided their time, usually on short notice.

The review process was similar to that used in previous years. Each paper was reviewed by at least three reviewers and then considered further by at least three ACs before a decision was made. Borderline papers and candidates for orals and spotlights were discussed in groups of three non-conflicted ACs. Oral and spotlight recommendations were then made by panels of ACs. Program Chairs did not submit papers, which allowed them to avoid conflicts throughout the process.

Past attendees know that as CVPR has grown it has become harder to see all of the amazing content one might hope to. It is tempting to try to rein in the growth of the conference program accordingly, but we feel strongly that maintaining the health of the research community is our highest priority. Paper decisions were made with the same high standards we have always had without trying to artificially cap the total number of accepted papers. This meant accepting 979 papers, 25% more than in 2017 and 52% more than in 2016. These include 70 papers selected for oral presentation (2.1%), 224 papers selected for spotlight oral presentation (6.6%), and 685 papers selected for poster presentation (20.4%). Overall, 29% of submitted papers were accepted for publication, with 30% of accepted papers selected for some form of oral presentation. As in recent years, all accepted papers appear in the poster sessions.

Prior to 2014, CVPR used a three-day, two-track format. As we've grown, the organizers of recent meetings have experimented with a variety of formats, including keeping the same format while including more papers (2015), adding a fourth day (2014, 2016), adding a fourth day with a half-day break during the meeting (2017), and adding a third parallel track (2017). In each case attendee feedback was gathered and used to guide future meetings. However, with our three-year (now four-year) planning window, some decisions had to be locked in years in advance. We made the decision back in 2015 for CVPR 2018 to be a three-day meeting, but the abundance of space in the facility makes it possible to again hold three parallel oral presentation tracks as well as to comfortably grow our space for posters alongside our industrial exhibition.

To accommodate both more attendees and more papers, we have experimented with a few changes this year. First, we made the decision to invite only one keynote speaker and to use the time for additional oral or poster sessions. When that speaker had to withdraw for family reasons, we went ahead and used that time similarly. As a result, each day of the main program includes one more poster or oral session than in previous years. Second, to continue to provide lunch to a growing number of attendees, we have extended the period during which lunch will be served and similarly extended the poster sessions immediately before and after. This overlap between extended lunch and extended poster sessions hopefully means greater flexibility and less crowding. Third, the exhibition will be open continuously from 10:00 a.m. to 6:30 p.m., allowing greater flexibility as well. Finally, we are

Message from the General and Program Chairs

experimenting with offering a few tutorials during the main program days in parallel with the paper sessions, providing more options for newer students and others attending the conference.

Our industrial exhibition also continues to grow, and we are grateful for the participation of our sponsors and other exhibitors. This exhibition provides a great opportunity for researchers, students, and other attendees to connect with companies ranging from promising startups to industry leaders. Exhibitor fees, plus generous donations from our corporate sponsors, allow us to improve the attendee experience while keeping CVPR the incredible value it has always been.

Two other changes are also worth noting: First, we have instituted a new "Code of Conduct", which can be found on the conference website. We trust that everyone involved in CVPR will continue to maintain the high degree of professionalism and collegiality we have always enjoyed here and hope this addition will be unnecessary, but we also feel it is important to be clear about our community expectations. Second, we have added a "Good Citizen of CVPR" panel on Friday for all who are new to our community and want to increase and improve their involvement.

Finally, we would like to thank everyone involved in making CVPR 2018 a success. This includes the organizing committee, the area chairs, the reviewers, emergency reviewers, authors, demo session participants, donors, exhibitors, and everyone else without whom this meeting would not be possible. We are grateful to Hall-Erickson International, Freeman, Presentation Technologies, and other vendors for their helpful coordination and experience. We are also grateful that Eric Mortensen, who has for many years managed the CVPR publication process, graciously agreed to reprise his role as publications chair one more time for us. A huge thank-you also goes to Nicole Finn, Liz Ryan, and C to C Events for their organization of the logistics of the conference.

Most of all, 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 your travel plans permit you time before or after the conference to get out from the city and to enjoy the rich natural diversity of the mountains, lakes, rivers, deserts, and National Parks of Utah. Enjoy CVPR 2018!

Program Chairs: **David Forsyth, Ivan Laptev, Aude Oliva, and Deva Ramanan**

General Chairs: **Michael S. Brown, Bryan Morse, and Shmuel Peleg**

CVPR 2018 Organizing Committee

General Chairs:	Michael S. Brown Bryan Morse Shmuel Peleg
Program Chairs:	David Forsyth Ivan Laptev Deva Ramanan Aude Oliva
Workshops Chairs:	Srikumar Ramalingam Mathieu Salzmann
Tutorials Chairs:	M. Pawan Kumar Andrea Vedaldi
Finance Chairs:	Walter Scheirer Ramin Zabih
Publications Chairs:	Eric Mortensen William Brendel
Demos Chair:	Seon Joo Kim
Website Chairs:	Dana Berman Abdelrahman Kamel
Corporate Relations Chairs:	Rogerio Feris Scott McCloskey Yu Wing Tai
Doctoral Consortium Chairs:	Peter Carr Xiaoming Liu
Student Volunteers Chairs:	Danna Gurari Marcus Brubaker
Local Arrangements Chair:	Ryan Farrell

Monday, June 18

0730–1830 Registration (South Lobby)

0730–0900 Breakfast (Hall A; Halls 1-4)

1000–1100 Morning Break (Hall A; Halls 1-4)

1200–1330 Lunch (Hall A; Halls 1-4)

1530–1630 Afternoon Break (Hall A; Halls 1-4)

1900–2000 PAMI Technical Committee Meeting (Ballroom)

A Crash Course on Human Vision

Organizers: Ali Borji
Krista A. Ehinger
James H. Elder
Odelia Schwartz
Thomas Serre

Location: Room 155 C

Time: 0845-1730 (Full Day)

Description: This tutorial provides a broad overview to the process of human visual perception, with an emphasis on how understanding the human visual system can help to advance computer vision research. This tutorial will cover both physiological and psychophysical approaches to understanding human vision and will relate the two fields together to create a consistent and complete understanding of the process of visual perception. For the physiological approach, the course will discuss the areas of lower-level visual processing in the receptors of the eye and the lateral geniculate nucleus and higher-level visual processing in different areas of the brain. In the psychophysical approach, the course will introduce the different psychophysical models of human vision, including the models of perceptual organization, perceptual segregation, and construction. Concepts of color, depth, movement and their visual perception will be introduced. To relate the materials presented in the context of different areas of computer science, examples of the quantification and use of these physiological and psychophysical models in computer vision, computer graphics, multimedia and HCI will be referenced.

Schedule:

0845 Introduction & Welcome
0900 Light and Image Formation on the Retina, *Ali Borji*
0945 Early Visual Processing, *Thomas Serre*

1030 Morning Break

1100 Color Processing, *Ali Borji*
1145 Motion Processing, *Krista A. Ehinger*

1230 Lunch

1400 Perceiving Depth and Size, *Krista A. Ehinger*
1445 Visual Attention, *John Tsotsos*

1530 Afternoon Break

1600 High-Level Processing, *Thomas Serre*
1645 Object Recognition, *Mazyar Fallah*

Interpreting and Explaining Deep Models in Computer Vision

Organizers: Wojciech Samek
Grégoire Montavon
Klaus-Robert Müller

Location: Room 355 E-F

Time: 0830-1200 (Half Day — Morning)

Description: Machine learning techniques such as deep neural networks (DNN) are able convert large amounts of data into highly predictive models. In complement to their unmatched predictive capability, it is becoming increasingly important to understand qualitatively and quantitatively how these models decide. Our tutorial will provide a broad overview of techniques for interpreting deep models, and how some of these techniques can be made useful on practical problems. In the first part we will lay a taxonomy of these methods, and explain how the various interpretation techniques can be characterized conceptually and mathematically. The second part of the tutorial will explain when and why we need interpretability.

Computer Vision for Robotics and Driving

Organizers: Anelia Angelova
Sanja Fidler

Location: Room 255 E-F

Time: 0830-1200 (Half Day — Morning)

Description: In this tutorial we will consider computer vision techniques that are key to robotics and autonomous driving. The main focus will be on novel cutting-edge deep learning approaches in these domains.

More specifically, the tutorial will start with a brief overview of Deep Learning approaches in the three major topics in computer vision: classification, detection and segmentation, then we will discuss concrete methods in computer vision applied to robotics and autonomous driving, such as object detection (in 2D and 3D), scene understanding, 3D analysis, learning from consecutive frames, learning with structure, learning with self-supervision, and others.

Schedule:

0830 Deep Learning for Computer Vision, *Anelia Angelova*
0915 Perception for Robotics, *Anelia Angelova*

1000 Afternoon Break

1030 Deep Learning for 3D, *Sanja Fidler*
1115 Vision for Driving, *Anelia Angelova*

Weakly Supervised Learning for Computer Vision

Organizers: Rodrigo Benenson
Hakan Bilen
Jasper Uijlings

Location: Room 155 D-F

Time: 0900-1215 (Half Day — Morning)

Description: Deep convolutional networks have become the go-to technique for a variety of computer vision task such as image classification, object detection, segmentation, key point detection, etc. These over-parameterized models are known to be data-hungry; tens of thousand of labelled examples are typically required. Since manual annotations are expensive, learning from "weaker" annotations (e.g. only image-level category labels to localize object instances by a bounding box) become key to expand the success of deep networks to new applications. This tutorial will provide an overview of weakly supervised learning methods in computer vision, and we will discuss the broad area of weakly supervised object recognition and its limitations of current state-of-the-art, evaluation metrics, and future ideas that will spur disruptive progress in the field of weakly supervised learning.

Schedule:

0900 Introduction, *Hakan Bilen (Univ. of Edinburgh)*
0915 Weakly Supervised Object Detection, *Hakan Bilen (Univ. of Edinburgh)*

1000 Morning Break

1030 Weakly Supervised Semantic Segmentation, *Rodrigo Benenson (Google)*
1115 Less Weakly Supervised Object Detection and Segmentation, *Jasper Uijlings (Google)*
1200 Closing Remarks, *Rodrigo Benenson (Google)*

Multi-View Visual Data Analytics

Organizers: Zhengming Ding
Ming Shao
Yun Fu

Location: Room 151 A-C

Time: 0830-1230 (Half Day — Morning)

Description: Multi-view data are extensively accessible nowadays thanks to various types of features, view-points and different sensors. For example, the most popular commercial depth sensor, Kinect, uses both visible light and near infrared sensors for depth estimation; automatic driving uses both visual and radar sensors to produce real-time 3D information on the road; and face analysis algorithms prefer face images from different views for high-fidelity reconstruction and recognition. All of them tend to facilitate better data representation in different application scenarios. Essentially, multiple features attempt to uncover various knowledge within each view to alleviate the final tasks, since each view would preserve both shared and private information. Recently there are several approaches proposed to deal with the multi-view visual data. Our tutorial covers most multi-view visual data representation

approaches, centered around several major applications, e.g., multi-view clustering, multi-view classification, and zero-shot learning. It discusses the current and upcoming challenges.

Schedule:

0830 Opening, *Zhengming Ding, Ming Shao*

0900 Unsupervised Multi-View Visual Data Analysis, *Ming Shao*

1030 Morning Break

1100 Supervised Multi-View Visual Data Analysis, *Zhengming Ding*

Motion Averaging: A Framework for Efficient and Accurate Large-Scale Camera Estimation in 3D Vision

Organizers: Venu Madhav Govindu

Location: Room 151 D-F

Time: 0830-1230 (Half Day — Morning)

Description: In recent years there has been growing interest in large-scale 3D reconstruction using both RGB and depth cameras. The concomitant need for accuracy, efficiency and scalability in camera motion estimation is addressed by the framework of motion averaging. Given many relative motion estimates between pairs of cameras, motion averaging solves for the 3D motions of individual cameras. The efficacy of motion averaging has attracted research interest leading to significant theoretical and algorithmic maturity. Owing to its major advantages over conventional approaches, motion averaging is now utilised in many 3D reconstruction pipelines. This tutorial provides a comprehensive introduction to motion averaging in 3D vision. An intuitive and systematic understanding of the underlying geometry of matrix Lie groups is developed. A comparative classification and summarization of various motion averaging methods is presented. In addition, this tutorial provides an exposition of algorithms and best practices. Along with developing a clear understanding of the state-of-the-art, this tutorial aims to enable researchers to utilise motion averaging principles in novel contexts of large-scale structure-from-motion as well as dense 3D modeling using depth cameras.

Interpretable Machine Learning for Computer Vision

Organizers: Bolei Zhou
Laurens van der Maaten
Been Kim
Andrea Vedaldi

Location: Room 355 E-F

Time: 1400-1735 (Half Day — Afternoon)

Description: Complex machine learning models such as deep convolutional neural networks and recursive neural networks have made great progress in a wide range of computer vision applications, such as object/scene recognition, image captioning, visual question answering. But they are often perceived as black-boxes. As models are going deeper in search of better recognition accuracy, it becomes even harder to understand the predictions given by the models and why. This tutorial is to broadly engage the computer vision community with the topic of interpretability and explainability in models used in computer vision. We will introduce the definition of interpretability and why it is important, and have a review on visualization and interpretation methodologies for analyzing both the data and the models in computer vision.

Schedule:

1400 Welcome & Overview

1410 Introduction to Interpretable Machine Learning, *Been Kim*

1450 Dos and Don'ts of Using t-SNE to Understand Vision Models, *Laurens van der Maaten*

1530 Afternoon Break

1615 Revisiting the Importance of Single Units in Deep Networks, *Bolei Zhou*

1655 Understanding Deep Networks Using Natural Pre-Images, Meaningful Perturbations, and Vector Embeddings, *Andrea Vedaldi*

Human Activity Recognition

Organizers: Michael S. Ryoo
Greg Mori
Kris Kitani

Location: Room 255 E-F

Time: 1330-1710 (Half Day — Afternoon)

Description: In the recent years, the field of human activity recognition has grown dramatically, reflecting its importance in many high-impact societal applications including smart surveillance, web-video search and retrieval, quality-of-life devices for elderly people, and robot perception. With the initial success of convolutional network models to learn video representations, the field is gradually moving towards detecting and forecasting more complex human activities involving multiple people, objects, and sub-events in various realistic scenarios. New important research topics and problems are appearing as a consequence, including (i) reliable spatio-temporal localization of activities, (ii) end-to-end modeling of activities' temporal structure and hierarchy, (iii) group activity recognition, (iv) activity forecasting, as well as (v) construction of large-scale datasets and convolutional models. The objective of this tutorial is to introduce and overview recent progress in these emerging topics, as well as to discuss, motivate and encourage future research in diverse subfields of activity recognition.

Schedule:

1330 Introduction
1350 Spatio-Temporal Activity Detection, *Greg Mori (SFU)*
1410 Learning Temporal Hierarchy, *Michael Ryoo (Indiana Univ.)*
1440 Invited Talk: Kinetics 600, *Joao Carreira (DeepMind)*
1505 Invited Talk: Observing Humans in their Natural Habitat: Datasets and Models, *Gunnar Sigurdsson (CMU)*

1530 Afternoon Break

1600 Group Activity Recognition, *Greg Mori (SFU)*
1620 Activity Forecasting, *Nick Rhinehart (CMU)*
1650 Emerging Topics (Including Privacy-Preservation), *Michael Ryoo (Indiana Univ.)*

Big Data Summarization: Algorithms and Applications

Organizers: Ehsan Elhamifar
Amit Roy-Chowdhury
Amin Karbasi

Location: Room 151 D-F

Time: 1345-1730 (Half Day — Afternoon)

Description: The increasing amounts of data in computer vision requires robust and scalable summarization tools to efficiently extract most important information from massive datasets. However, summarization involves optimization programs that are nonconvex and NP-hard, in general. While convex, nonconvex and submodular optimization have been studied intensively in mathematics, successful and effective applications of them for information summarization along with new theoretical results have recently emerged. These results, in contrast with more classical approaches, can deal with structured data, nonlinear models, data nuisances and exponentially large dataset. The goal of this tutorial is to present the audience with a unifying perspective of this problem, introducing the basic concepts and connecting nonconvex methods with convex sparse optimization as well as submodular optimization. The presentation of the formulations, algorithms and theoretical foundations will be complemented with applications in computer vision, including video and image summarization, procedure learning from instructional data, pose estimation, active learning and more.

Schedule:

1345 Overview of Summarization Algorithms: Modeling, Optimizations, Applications, *Ehsan Elhamifar*
1430 Submodular Optimization Methods for Summarization, *Amin Karbasi*

1530 Afternoon Break

1600 Sequential Data Summarization and Applications to Procedure Learning, *Ehsan Elhamifar*
1645 Collaborative Summarization With Side Information, *Amit Roy-Chowdhury*

Optimisation in Multiple View Geometry: The L-Infinity Way

Organizers: Tat-Jun Chin
Anders Eriksson
Fredrik Kahl

Location: Room 151 A-C

Time: 1330-1650 (Half Day — Afternoon)

Description: Stemming from its roots in photogrammetry, the theory and practice of multiple view geometry has traditionally focussed on L2 minimisation. However, due to the projective nature of imaging, many of the optimisation problems associated with L2 minimisation are plagued by multiple local minima. This prevents globally optimal solutions and necessitates careful preprocessing and/or initialisation to ensure success. The situation changes dramatically, however, if we perform L-infinity minimisation, where it can be shown that many geometric problems are amenable to global solutions. There has been significant progress in the "L-infinity way", encompassing aspects such as robust estimation, large-scale optimisation, and guaranteed approximations. This tutorial aims to give an in-depth introduction of L-infinity optimisation in geometric vision. We emphasise on the basic mathematical and algorithmic concepts, so as to convey deeper understanding and appreciation of the approach. We also survey important recent advances in the area, which contribute towards making the L-infinity way a bona fide alternative to current methods for geometric vision.

Schedule:

1330 Arrival & Welcome, *Tat-Jun Chin*
1340 Introduction and Core Algorithms, *Fredrik Kahl*
1430 Approximate Algorithms and Robust Estimation, *Tat-Jun Chin*

1520 Afternoon Break

1600 Large-Scale Optimisation and Applications, *Anders Eriksson*

Unsupervised Visual Learning

Organizers: Pierre Sermanet
Carl Vondrick
Anelia Angelova

Location: Room 155 D-F

Time: 1400-1800 (Half Day — Afternoon)

Description: Unsupervised learning focuses on learning from vast amounts of data without manual supervision. While the topic of unsupervised learning is very rich, and with extensive prior research, we here focus on unsupervised learning for visual data. More specifically we will focus on recent approaches for building visual representation learning e.g. learning from context, learning from video and learning from future frames. We also consider powerful methods emerging from self-supervision, e.g. learning across sensors, or exploiting feature correlations in video, or from different viewpoints. Some of these powerful methods emerging from self-supervision are already finding applications in computer vision and robotics.

Schedule:

1400 Unsupervised Learning for Vision, *Anelia Angelova*
1430 Representation Learning from Unlabeled Images and Videos, *Carl Vondrick*

1530 Afternoon Break

1600 Unsupervised Learning From 3D, *Anelia Angelova*
1630 Self-Supervised Learning for Vision and Robotics, *Pierre Sermanet*

Notes:

Monday, June 18

0730–1830 Registration (South Lobby)

0730–0900 Breakfast (Hall A; Halls 1-4)

1000–1100 Morning Break (Hall A; Halls 1-4)

1200–1330 Lunch (Hall A; Halls 1-4)

1530–1630 Afternoon Break (Hall A; Halls 1-4)

1900–2000 PAMI Technical Committee Meeting (Ballroom)

Low-Power Image Recognition Challenge

Organizers: Yung-Hsiang Lu
Alexander C. Berg
Yiran Chen
Bo Chen
Yang Lu

Location: Room 250 A-B

Schedule: Full Day

0900 **Welcome:** *Terence Martinez (IEEE)*

0910 **Track 1 & Track 2 Winners**

0920 **Invited Talk:** A Systematic Approach to Benchmark ML Workload on Server, Mobile and Embedded Platforms, *Fei Sun (Facebook)*

0950 **Invited Talk:** Accelerating Research in Mobile Computer Vision Through Common and Accessible On-Device Benchmarks, *Bo Chen (Google)*

1010 **Invited Talk:** Best Practices in Developing Mobile Computer Vision Models, *Andrew Howard (Google)*

1030 Morning Break (Hall A; Halls 1-4)

1040 **Track 3 Competition:** Schedule TBA

Disguised Faces in the Wild

Organizers: Rama Chellappa
Nalini Ratha
Mayank Vatsa
Richa Singh

Location: Room 251 A

Schedule: Full Day

0820 **Opening Remarks**

S1: Oral Session 1 (0830-0930)

0830 Deep Features for Recognizing Disguised Faces in the Wild, *Ankan Bansal; Rajeev Ranjan; Carlos D. Castillo; Rama Chellappa*

0850 Face Verification With Disguise Variations via Deep Disguise Recognizer, *Naman Kohli; Daksha Yadav; Afzel Noore*

0910 DisguiseNet: A Contrastive Approach for Disguised Face Verification in the Wild, *Skand Vishwanath Peri; Abhinav Dhall*

0930 **Invited Talk:** *Christopher Boehnen (IARPA)*

1030 Morning Break (Hall A; Halls 1-4)

S2: Oral Session 2 (1100-1200)

1100 Deep Disguised Faces Recognition, *Kaipeng Zhang; Ya-Liang Chang; Winston Hsu*

1120 Hard Example Mining With Auxiliary Embeddings, *Evgeny Smirnov; Aleksandr Melnikov; Andrei Oleinik; Elizaveta Ivanova; Ilya Kalinovskiy; Eugene Luckyanets*

1140 Detecting Presentation Attacks From 3D Face Masks Under Multispectral Imaging, *Jun Liu; Ajay Kumar*

1200 Lunch (Hall A; Halls 1-4)

S3: Phase 2 Submission Presentations (1330-1430)

1330 **Oral Presentation:** Phase 2

1350 **Oral Presentation:** Phase 2

1410 **Oral Presentation:** Phase 2

1430 **Keynote Talk:** *Richard W. Vorder Bruegge (FBI)*

1530 **Summary:** Disguised Faces in the Wild, *Vineet Kushwaha, Maneet Singh, Richa Singh, Mayank Vatsa, Nalini Ratha, Rama Chellapa*

1545 **Awards**

NVIDIA AI City Challenge

Organizers: Milind Naphade
Rama Chellappa
Jenq-Neng Hwang
Ming-Ching Chang
Ming-Yu Liu
Siwei Lyu
David C. Anastasiu
Anuj Sharma
Zeyu Gao

Location: Room 355 B

Schedule: Full Day

0900 **Opening Remarks** (Challenge Summary)

0925 **Keynote Talk:** Video Analytics in Traffic and Public Safety, *John Garofolo (NIST)*

S1: Oral Presentations I (0950-1030)

0950 Video Analytics in Smart Transportation for the AIC'18 Challenge, *Ming-Ching Chang; Yi Wei; Nenghui Song; Siwei Lyu*

1010 Challenges on Large Scale Surveillance Video Analysis, *Weitao Feng; Deyi Ji; Yiru Wang; Shuorong Chang; Hansheng Ren; Weihao Gan*

1030 Morning Break (Hall A; Halls 1-4)

S2: Oral Presentations II (1100-1220)

1100 Graph@FIT Submission to the NVIDIA AI City Challenge 2018, *Jakub Sochor; Jakub Špaňhel; Roman Juránek; Petr Dobeš; Adam Herout*

1120 AIC2018 Report: Traffic Surveillance Research, *Tingyu Mao; Wei Zhang; Haoyu He; Yanjun Lin; Vinay Kale; Alexander Stein; Zoran Kostic*

1140 Speed Estimation and Abnormality Detection From Surveillance Cameras, *Panagiotis Giannakeris; Vagia Kaltsa; Konstantinos Avgerinakis; Alexia Briassouli; Stefanos Vrochidis; Ioannis Kompatsiaris*

1200 Traffic Flow Analysis With Multiple Adaptive Vehicle Detectors and Velocity Estimation With Landmark-Based Scanlines, *Minh-Triet Tran; Tung Dinh-Duy; Thanh-Dat Truong; Vinh Ton-That; Thanh-Nhon Do; Quoc-An Luong; Thanh-An Nguyen; Vinh-Tiep Nguyen; Minh N. Do*

1220 Lunch (Hall A; Halls 1-4)

S3: Oral Presentations III (1320-1520)

1320 Single-Camera and Inter-Camera Vehicle Tracking and 3D Speed Estimation Based on Fusion of Visual and Semantic Features, *Zheng Tang; Gaoang Wang; Hao Xiao; Aotian Zheng; Jenq-Neng Hwang*

1340 Geometry-Aware Traffic Flow Analysis by Detection and Tracking, *Honghui Shi*

1400 Vehicle Re-Identification With the Space-Time Prior, *Chih-Wei Wu; Chih-Ting Liu; Cheng-En Chiang; Wei-Chih Tu; Shao-Yi Chien*

1420 Unsupervised Anomaly Detection for Traffic Surveillance Based on Background Modeling, *JiaYi Wei; JianFei Zhao; YanYun Zhao; ZhiCheng Zhao*

1440 A Semi-Automatic 2D Solution for Vehicle Speed Estimation From Monocular Videos, *Amit Kumar; Pirazh Khorramshahi; Wei-An Lin; Prithviraj Dhar; Jun-Cheng Chen; Rama Chellappa*

1500 Dual-Mode Vehicle Motion Pattern Learning for High Performance Road Traffic Anomaly Detection, *Yan Xu; Xi Ouyang; Yu Cheng; Shining Yu; Lin Xiong; Choon-Ching Ng; Sugiri Pranata; Shengmei Shen; Junliang Xing*

1520 **Panel:** Video Analytics in Traffic and Public Safety

- **Moderator:** *Milind Naphade (NVIDIA Corporation)*
- **Panelists:** *Terry Adams (I-ARPA); David Ness (City of Dubuque, Traffic Engineer); David Kuehn (FHWA, USDOT); Anuj Sharma (Iowa State Univ.); John Garofolo (NIST)*

1600 Afternoon Break (Hall A; Halls 1-4)

S4: Poster Presentation (1600-1700)

- Vehicle Tracking and Speed Estimation From Traffic Videos, *Shuai Hua; Manika Kapoor; David C. Anastasiu*
- Traffic Speed Estimation From Surveillance Video Data, *Tingting Huang*
- Unsupervised Vehicle Re-Identification Using Triplet Networks, *Pedro Antonio Marín-Reyes; Andrea Palazzi; Luca Bergamini; Simone Calderara; Javier Lorenzo-Navarro; Rita Cucchiara*

1700 **Awards Ceremony**

Robust Vision Challenge

Organizers: Andreas Geiger
Matthias Niessner
Marc Pollefeys
Carsten Rother
Daniel Scharstein
Hassan Alhaija
Angela Dai
Katrin Honauer
Joel Janai
Torsten Sattler
Nick Schneider
Johannes Schoenberger
Thomas Schoeps
Jonas Uhrig
Jonas Wulff
Oliver Zendel

Location: Room 355 C

Schedule: Full Day

0845 **Welcome**
0900 **Invited Talk:** *Judy Hoffman (UC Berkeley)*
0945 **Introduction & Winner Talks:** Stereo Challenge

1030 Morning Break (Hall A; Halls 1-4)

1045 **Introduction & Winner Talks:** Multiview Stereo Challenge
1130 **Introduction & Winner Talks:** Optical Flow Challenge

1215 Lunch (Hall A; Halls 1-4)

1330 **Invited Talk:** *Uwe Franke (Daimler AG)*
1415 **Introduction & Winner Talks:** Single Image Depth Prediction Challenge
1500 **Introduction & Winner Talks:** Semantic Segmentation Challenge

1545 Afternoon Break (Hall A; Halls 1-4)

1600 **Invited Talk:** *Dawn Song (UC Berkeley)*
1645 **Introduction & Winner Talks:** Semantic Instance Segmentation Challenge
1730 **Discussion & Closing Remarks**
1900 **Dinner**

VQA Challenge and Visual Dialog

Organizers: Yash Goyal
Satwik Kottur
Abhishek Das
Aishwarya Agrawal
Stefan Lee
Dhruv Batra
Devi Parikh

Location: Room 155 A

Schedule: Full Day

0900 **Welcome**
0910 **Invited Talk:** *Jeffrey Bigham (Carnegie Mellon Univ.)*
0935 **Invited Talk:** *Adriana Kovashka (Univ. of Pittsburgh)*
1000 **Invited Talk:** *Jacob Andreas (Univ. of California at Berkeley)*

1025 Morning Break (Hall A; Halls 1-4)

1050 **Invited Talk:** *Jitendra Malik (Facebook AI Research & Univ. of California at Berkeley)*
1115 **Overview of Dataset, Challenge, Winner Announcements, Analysis of Results**
1145 **Challenge Talk:** Honorable Mention
1150 **Challenge Talk:** Runner-Up
1205 **Challenge Talk:** Winner

1220 Lunch (Hall A; Halls 1-4)

1345 **Invited Talk:** *Ross Girshick (Facebook AI Research)*
1410 **Invited Talk:** *Nasrin Mostafazadeh (BenevolentAI)*

1435 Afternoon Break (Hall A; Halls 1-4) **& Poster Session**

1550 **Invited Talk:** *Aaron Courville (Univ. of Montreal)*
1615 **Invited Talk:** *Peter Anderson (Australian National Univ.)*
1640 **Panel:** Future Directions
1725 **Closing Remarks**

Deep Learning for Visual SLAM

Organizers: Ronald Clark
Sudeep Pillai
Alex Kendall
Will Maddern
Andrew Davison
Stefan Leutenegger

Location: Room 255 C

Schedule: Full Day

0830 **Welcome**

0845 **Invited Talk:** *Jitendra Malik (Berkley)*

0915 **Invited Talk:** *Simon Lucey (CMU)*

0945 SuperPoint: Self-Supervised Interest Point Detection and Description, *Daniel DeTone; Tomasz Malisiewicz; Andrew Rabinovich*

1000 Morning Break (Hall A; Halls 1-4)

1030 **Invited Talk:** *Jan Kautz (NVIDIA)*

1100 **Invited Talk:** *Daniel Cremers (TUM)*

1130 **Invited Talk:** *Katerina Frakiadaki (CMU)*

1200 Global Pose Estimation With an Attention-Based Recurrent Network, *Emilio Parisotto; Devendra Singh Chaplot; Jian Zhang; Ruslan Salakhutdinov*

1235 Lunch (Hall A; Halls 1-4) **& Poster Session**

- Visual SLAM for Automated Driving: Exploring the Applications of Deep Learning, *Stefan Milz; Georg Arbeiter; Christian Witt; Bassam Abdallah; Senthil Yogamani*
- Mask-SLAM: Robust Feature-Based Monocular SLAM by Masking Using Semantic Segmentation, *Masaya Kaneko; Kazuya Iwami; Toru Ogawa; Toshihiko Yamasaki; Kiyoharu Aizawa*
- Geometric Consistency for Self-Supervised End-to-End Visual Odometry, *Ganesh Iyer; J. Krishna Murthy; Gunshi Gupta; Madhava Krishna; Liam Paull*
- Learning Descriptor, Confidence, and Depth Estimation in Multi-View Stereo, *Sungil Choi; Seungryong Kim; Kihong Park; Kwanghoon Sohn*
- DepthNet: A Recurrent Neural Network Architecture for Monocular Depth Prediction, *Arun CS Kumar; Suchendra M. Bhandarkar; Mukta Prasad*
- Towards CNN Map Representation and Compression for Camera Relocalisation, *Luis Contreras; Walterio Mayol-Cuevas*
- Monocular Depth Prediction Using Generative Adversarial Networks, *Arun CS Kumar; Suchendra M. Bhandarkar; Mukta Prasad*
- Learning 3D Scene Semantics and Structure From a Single Depth Image, *Bo Yang; Zihang Lai; Xiaoxuan Lu; Shuyu Lin; Hongkai Wen; Andrew Markham; Niki Trigoni*
- QuadricSLAM: Dual Quadrics As SLAM Landmarks, *Lachlan Nicholson; Michael Milford; Niko Sünderhauf*

1435 **Invited Talk:** *Noah Snavely (Google)*

1500 **Invited Talk:** *Vladlen Koltun (Intel)*

1530 **Invited Talk:** *Matthias Niessner (TUM)*

1630 **Panel:** Challenges, Potentials and the Future of Machine Learning and SLAM, *All speakers, Chair: Andrew Davison (Imperial College London)*

1730 **Dinner**

DeepGlobe: A Challenge for Parsing the Earth through Satellite Images

Organizers: Ilke Demir
Manohar Paluri
Lorenzo Torresani
Ramesh Raskar
Daniel Aliaga
Nikhil Naik
Kris Koperski
Raffay Hamid

Location: Room 150 G

Schedule: Full Day

0900 **Welcome**

S1: Road Extraction for Maps and Beyond (0915-1035)

0915 **Invited Talk:** Geospatial AI for Impact, *Ramesh Raskar (MIT)*

0945 **Invited Talk:** Roads and Maps, *Raquel Urtasun (Univ. of Toronto; Uber ATG)*

1015 **Challenge Winner Presentation —** Road Extraction

1035 Morning Break (Hall A; Halls 1-4)

S2: Building Detection for Population Dynamics (1045-1205)

1045 **Invited Talk:** Mapping Economic Development from Space, *Stefano Ermon (Stanford Univ.)*

1115 **Invited Talk:** AI in Global Scale Mapping and 3D Modeling, *HakJae Kim (IARPA)*

1145 **Challenge Winner Presentation —** Building Detection

1205 Lunch (Hall A; Halls 1-4)

S3: Land Cover Classification for Innovation and Impact (1300-1420)

1300 **Invited Talk:** Lands and Innovation, *Nebojsa Jojic (Microsoft Research), Mary Wahl (Microsoft Research)*

1330 **Invited Talk:** Creating Large Scale Datasets for Object Detection in Satellite Imagery, *Gopal Erinjippurath (Planet Labs)*

1400 **Challenge Winner Presentation —** Land Cover Classification

1420 **Poster Session**

- Semantic Binary Segmentation Using Convolutional Networks Without Decoders, *Shubhra Aich; William van der Kamp; Ian Stavness*
- Stacked U-Nets With Multi-Output for Road Extraction, *Tao Sun; Zehui Chen; Wenxiang Yang; Yin Wang*
- D-LinkNet: LinkNet With Pretrained Encoder and Dilated Convolution for High Resolution Satellite Imagery Road Extraction, *Lichen Zhou; Chuang Zhang; Ming Wu*
- Fully Convolutional Network for Automatic Road Extraction From Satellite Imagery, *Alexander Buslaev; Selim Seferbekov; Vladimir Iglovikov; Alexey Shvets*
- Road Detection With EOSResUNet and Post Vectorizing Algorithm, *Oleksandr Filin; Anton Zapara; Serhii Panchenko*
- Residual Inception Skip Network for Binary Segmentation, *Jigar Doshi*
- Roadmap Generation Using a Multi-Stage Ensemble of Deep Neural Networks With Smoothing-Based Optimization, *Dragoş Costea; Alina Marcu; Emil Sluşanschi; Marius Leordeanu*
- Rotated Rectangles for Symbolized Building Footprint Extraction, *Matt Dickenson; Lionel Gueguen*
- Building Detection From Satellite Imagery Using a Composite Loss Function, *Sergey Golovanov; Rauf Kurbanov; Aleksey Artamonov; Alex Davydow; Sergey Nikolenko*
- Building Detection From Satellite Imagery Using Ensemble of Size-Specific Detectors, *Ryuhei Hamaguchi; Shuhei Hikosaka*
- TernausNetV2: Fully Convolutional Network for Instance Segmentation, *Vladimir Iglovikov; Selim Seferbekov; Alexander Buslaev; Alexey Shvets*
- Semantic Segmentation Based Building Extraction Method Using Multi-Source GIS Map Datasets and Satellite Imagery, *Weijia Li; Conghui He; Jiarui Fang; Haohuan Fu*
- CNNs Fusion for Building Detection in Aerial Images for the Building Detection Challenge, *Remi Delassus; Romain Giot*
- Building Extraction From Satellite Images Using Mask R-CNN With Building Boundary Regularization, *Kang Zhao; Jungwon Kang; Jaewook Jung; Gunho Sohn*
- Deep Aggregation Net for Land Cover Classification, *Tzu-Sheng Kuo; Keng-Sen Tseng; Jia-Wei Yan; Yen-Cheng Liu; Yu-Chiang Frank Wang*
- Stacked U-Nets for Ground Material Segmentation in Remote Sensing Imagery, *Arthita Ghosh; Max Ehrlich; Sohil Shah; Larry S. Davis; Rama Chellappa*
- Land Cover Classification From Satellite Imagery With U-Net and Lovász-Softmax Loss, *Alexander Rakhlin; Alex Davydow; Sergey Nikolenko*
- Dense Fusion Classmate Network for Land Cover Classification, *Chao Tian; Cong Li; Jianping Shi*
- NU-Net: Deep Residual Wide Field of View Convolutional Neural Network for Semantic Segmentation, *Mohamed Samy; Karim Amer; Kareem Eissa; Mahmoud Shaker; Mohamed ElHelw*
- Feature Pyramid Network for Multi-Class Land Segmentation, *Selim Seferbekov; Vladimir Iglovikov; Alexander Buslaev; Alexey Shvets*
- Uncertainty Gated Network for Land Cover Segmentation, *Guillem Pascual; Santi Seguí; Jordi Vitrià*
- Land Cover Classification With Superpixels and Jaccard Index Post-Optimization, *Alex Davydow; Sergey Nikolenko*

1600 **Award Ceremony**

1630 **Discussion & Collaboration Session**, *Invited Speakers & Challenge Winners*

1730 **Closing Remarks**

1900 **Workshop Dinner:** *All Presenters & Workshop Team*

Visual Understanding of Humans in Crowd Scene and Look Into Person Challenge

Organizers: Xiaodan Liang
Jian Zhao
Jianshu Li
Liang Lin
Jiashi Feng
Eric Xing

Location: Room 250 D-E

Schedule: Full Day

0830 **Welcome & Opening Remarks**

0840 **Introductions & Results:** The Look Into Person (LIP) Challenge

0910 **Challenge Winner Talk:** Single-Person Human Parsing

0925 **Invited Talk:** *Xiansheng Hua (Alibaba Group)*

0955 Morning Break (Hall A; Halls 1-4) **& Poster Session**

- Adaptation and Re-Identification Network: An Unsupervised Deep Transfer Learning Approach to Person Re-Identification, *Yu-Jhe Li; Fu-En Yang; Yen-Cheng Liu; Yu-Ying Yeh; Xiaofei Du; Yu-Chiang Frank Wang*
- Attention in Multimodal Neural Networks for Person Re-Identification, *Aske R. Lejbølle; Benjamin Krogh; Kamal Nasrollahi; Thomas B. Moeslund*
- Pose Encoding for Robust Skeleton-Based Action Recognition, *Girum G. Demisse; Konstantinos Papadopoulos; Djamila Aouada; Björn Ottersten*
- An Aggregated Multicolumn Dilated Convolution Network for Perspective-Free Counting, *Diptodip Deb; Jonathan Ventura*
- Learning to Refine Human Pose Estimation, *Mihai Fieraru; Anna Khoreva; Leonid Pishchulin; Bernt Schiele*
- Crowd Activity Change Point Detection in Videos via Graph Stream Mining, *Meng Yang; Lida Rashidi; Sutharshan Rajasegarar; Christopher Leckie; Aravinda S. Rao; Marimuthu Palaniswami*

1025 **Invited Talk:** *Shuicheng Yan (Qihoo/360)*

1055 **Challenge Winner Talk:** Single-Person Pose Estimation

1110 **Challenge Winner Talk:** Multi-Person Human Parsing

1125 Lunch (Hall A; Halls 1-4)

1400 **Invited Talk:** *Trevor Darrell (UC Berkeley)*

1430 **Challenge Winner Talk:** Fine-Grained Multiple-Human Parsing

1445 **Challenge Winner Talk:** Multiple-Human Pose Estimation

1500 Afternoon Break (Hall A; Halls 1-4) **& Poster session**

1530 **Invited Talk:** *Alan Yuille (Johns Hopkins Univ.)*

1600 **Awards & Future Plans**

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: Room 150 D-F

Schedule: Full Day

0830 **Opening Remarks**

0840 **Invited Talk:** Human-Machine Collaboration for Large-Scale Image Annotation, *Vittorio Ferrari (Google Research & Univ of Edinburgh)*

0920 **Dataset Update & Challenge Overview**

1000 Morning Break (Hall A; Halls 1-4)

1030 **Oral Presentation:** Invited Participant

1050 **Oral Presentation:** Invited Participant

1110 Lunch (Hall A; Halls 1-4) **& Poster Session**

1400 **Invited Talk:** Learning Single-Image 3D From the Web, *Jia Deng (Univ of Michigan)*

1440 **Invited Talk:** Learning From Web Data and Adapting Beyond It, *Boqing Gong (Tencent AI Lab; ICSI, UC Berkeley)*

1520 **Oral Presentation:** Invited Participant

1540 **Award Session & Closing Remarks**

1730 **Dinner**

Diff-CVML: Differential Geometry in Computer Vision and Machine Learning

Organizers: Anuj Srivastava
Baba Vemuri
Vittorio Murino
Rama Chellappa
Pavan Turaga
René Vidal
Richard Hartley

Location: Room 250 F

Schedule: Full Day

0900 **Welcome Remarks**

0905 **Poster Spotlights** (2 minutes each)

0930 **Keynote Talk:** Geometric Deep Learning on Manifolds and Graphs, *Michael Bronstein (Univ. of Lugano)*

1010 Morning Break (Hall A; Halls 1-4) **& Posters**

1030 The Riemannian Geometry of Deep Generative Models, *Hang Shao; Abhishek Kumar; P. Thomas Fletcher*

1050 Elastic Handling of Predictor Phase in Functional Regression Models, *Kyungmin Ahn; J. Derek Tucker; Wei Wu; Anuj Srivastava*

1110 Geodesic Discriminant Analysis for Manifold-Valued Data, *Maxime Louis; Benjamin Charlier; Stanley Durrleman*

1130 A Mixture Model for Aggregation of Multiple Pre-Trained Weak Classifiers, *Rudrasis Chakraborty; Chun-Hao Yang; Baba C. Vemuri*

1150 Temporal Alignment Improves Feature Quality: An Experiment on Activity Recognition With Accelerometer Data, *Hongjun Choi; Qiao Wang; Meynard Toledo; Pavan Turaga; Matthew Buman; Anuj Srivastava*

1210 Lunch (Hall A; Halls 1-4)

1400 **Keynote Talk:** The Geometry of the Loss Function for Deep Networks and Its Role in Generalization, *Stefano Soatto (Univ. of California, Los Angeles)*

1440 **Keynote Talk:** Optimisation Geometry, *Jonathan Manton (Univ. of Melbourne)*

1520 **Keynote Talk:** Probabilistic Geodesic Models for Dimensionality Reduction and Regression on Manifolds, *Tom Fletcher (Univ. of Utah)*

1600 **Keynote Talk:** Machine Learning Approaches for Medical Image Registration, *Marc Niethammer (Univ. of North Carolina, Chapel Hill)*

1640 **Poster Session**

- Locally-Weighted Elastic Comparison of Planar Shapes, *Justin Strait; Sebastian Kurtek; Steven MacEachern*
- Covariance Pooling for Facial Expression Recognition, *Dinesh Acharya; Zhiwu Huang; Danda Pani Paudel; Luc Van Gool*
- Image Segmentation by Deep Learning of Disjunctive Normal Shape Model Shape Representation, *Mehran Javanmardi; Ricardo Bigolin Lanfredi; Mujdat Cetin; Tolga Tasdizen*
- Predicting Dynamical Evolution of Human Activities From a Single Image, *Suhas Lohit; Ankan Bansal; Nitesh Shroff; Jaishanker Pillai; Pavan Turaga; Rama Chellappa*
- Covariance Matrices Encoding Based on the Log-Euclidean and Affine Invariant Riemannian Metrics, *Ioana Ilea; Lionel Bombrun; Salem Said; Yannick Berthoumieu*
- Principal Curvature Guided Surface Geometry Aware Global Shape Representation, *Somenath Das; Suchendra M. Bhandarkar*
- SphereNet: Learning Spherical Representations for Classification of Omnidirectional Images, *Benjamin Coors; Alexandru Condurache; Andreas Geiger*
- Minimal-Entropy Correlation Alignment for Unsupervised Deep Domain Adaptation, *Pietro Morerio; Jacopo Cavazza; Vittorio Murino*
- DEFRAG: Deep Euclidean Feature Representations Through Adaptation on the Grassmann Manifold, *Breton Minnehan; Andreas Savakis*
- Localizing Differentially Evolving Covariance Structures via Scan Statistics, *Ronak Mehta; Hyunwoo J. Kim; Shulei Wang; Sterling C. Johnson; Ming Yuan; Vikas Singh*

Notes:

Biometrics

Organizers: Bir Bhanu
Ajay Kumar

Location: Room 260

Schedule: Full Day

0815 **Opening Remarks**

S1: Face & Signature Verification (0820-0920)

0820 Toward More Realistic Face Recognition Evaluation Protocols for the YouTube Faces Database, *Yoanna Martínez-Díaz; Heydi Méndez-Vázquez; Leyanis López-Avila; Leonardo Chang; L. Enrique Sucar; Massimo Tistarelli*

0840 Dict Layer: A Structured Dictionary Layer, *Yefei Chen; Jianbo Su*

0900 Hierarchical Dictionary Learning and Sparse Coding for Static Signature Verification, *Elias N. Zois; Marianna Papagiannopoulou; Dimitrios Tsourounis; George Economou*

S2: Iris Recognition & Security (0920-1000)

0920 Realtime Quality Assessment of Iris Biometrics Under Visible Light, *Mohsen Jenadeleh; Marius Pedersen; Dietmar Saupe*

0930 Multi-Frame Super Resolution for Ocular Biometrics, *Narsi Reddy; Dewan Fahim Noor; Zhu Li; Reza Derakhshani*

0940 Face Template Protection Using Deep Convolutional Neural Network, *Arun Kumar Jindal; Srinivas Chalamala; Santosh Kumar Jami*

0950 Incorporating Touch Biometrics to Mobile One-Time Passwords: Exploration of Digits, *Ruben Tolosana; Rubén Vera-Rodríguez; Julian Fierrez; Javier Ortega-Garcia*

1000 Morning Break (Hall A; Halls 1-4)

1030 **Invited Talk:** Face Recognition by Deep Learning - The Imbalance Problem, *Chen Change Loy (The Chinese Univ. of Hong Kong)*

S3: Cross Domain & Low Resolution Face Recognition (1130-1230)

1130 Identity Aware Synthesis for Cross Resolution Face Recognition, *Maneet Singh; Shruti Nagpal; Mayank Vatsa; Richa Singh; Angshul Majumdar*

1150 GenLR-Net: Deep Framework for Very Low Resolution Face and Object Recognition With Generalization to Unseen Categories, *Sivaram Prasad Mudunuri; Soubhik Sanyal; Soma Biswas*

1210 Attribute-Centered Loss for Soft-Biometrics Guided Face Sketch-Photo Recognition, *Hadi Kazemi; Sobhan Soleymani; Ali Dabouei; Mehdi Iranmanesh; Nasser M. Nasrabadi*

S4: Fingerprint & Signature Verification (1230-1300)

1230 Latent Fingerprint Image Quality Assessment Using Deep Learning, *Jude Ezeobiejesi; Bir Bhanu*

1240 Unconstrained Fingerphoto Database, *Shaan Chopra; Aakarsh Malhotra; Mayank Vatsa; Richa Singh*

1250 Hybrid User-Independent and User-Dependent Offline Signature Verification With a Two-Channel CNN, *Mustafa Berkay Yılmaz; Kağan Öztürk*

1300 Lunch (Hall A; Halls 1-4)

S5: Face Detection & Recognition (1400-1430)

1400 It Takes Two to Tango: Cascading Off-the-Shelf Face Detectors, *Siqi Yang; Arnold Wiliem; Brian C. Lovell*

1410 Time Analysis of Pulse-Based Face Anti-Spoofing in Visible and NIR, *Javier Hernandez-Ortega; Julian Fierrez; Aythami Morales; Pedro Tome*

1420 A Deep Face Identification Network Enhanced by Facial Attributes Prediction, *Fariborz Taherkhani; Nasser M. Nasrabadi; Jeremy Dawson*

S6: Soft Biometrics & Presentation Attack Detection (1430-1530)

1430 Gait Recognition by Deformable Registration, *Yasushi Makihara; Daisuke Adachi; Chi Xu; Yasushi Yagi*

1450 Fusion of Handcrafted and Deep Learning Features for Large-Scale Multiple Iris Presentation Attack Detection, *Daksha Yadav; Naman Kohli; Akshay Agarwal; Mayank Vatsa; Richa Singh; Afzel Noore*

1510 Hierarchical Network for Facial Palsy Detection, *Gee-Sern Jison Hsu; Wen-Fong Huang; Jiunn-Horng Kang*

1530 Afternoon Break (Hall A; Halls 1-4)

1600 **Panel:** Impact of Deep Learning on Biometrics and Trends, *Chair: Brian Lovell (Univ. of Queensland)*

1700 **Awards, Valedictory, & Closing Remarks**

Embedded Vision

Organizers: Ravi Kumar Satzoda
Zoran Nikolic
Tse-Wei Chen
Rajesh Narasimha
Martin Humenberger

Location: Room 355 A

Schedule: Full Day

0830 **Opening Remarks**

S1: Oral 1 (0840-1000)

0840 **Invited Talk:** Vision for Autonomous Driving and Multitasking Humans, *Mohan Trivedi (UCSD)*

0920 A Comparative Study of Real-Time Semantic Segmentation for Autonomous Driving, *Mennatullah Siam; Mostafa Gamal; Moemen Abdel-Razek; Senthil Yogamani; Martin Jagersand; Hong Zhang*

0940 Efficient Semantic Segmentation Using Gradual Grouping, *Nikitha Vallurupalli; Sriharsha Annamaneni; Girish Varma; C.V. Jawahar; Manu Mathew; Soyeb Nagori*

1000 Morning Break (Hall A; Halls 1-4)

S2: Oral 2 (1030-1210)

1030 **Invited Talk:** Edge Computing With the Intel Neural Stick, *Cormac Brick (Intel)*

1110 IFQ-Net: Integrated Fixed-Point Quantization Networks for Embedded Vision, *Hongxing Gao; Wei Tao; Dongchao Wen; Tse-Wei Chen; Kinya Osa; Kato Masami*

1130 **Invited Talk:** Intelligent Scene Perception Onboard Autonomous Platforms, *Raghuveer Rao (US Army Research Lab)*

1210 Lunch (Hall A; Halls 1-4)

S3: Oral 3 (1310-1440)

1310 **Invited Talk:** Scalable and Semantic Indoor Mapping, *Donghwan Lee (Naver Labs)*

1400 Interpolation-Based Object Detection Using Motion Vectors for Embedded Real-Time Tracking Systems, *Takayuki Ujiie; Masayuki Hiromoto; Takashi Sato*

1420 Onboard Stereo Vision for Drone Pursuit or Sense and Avoid, *Cevahir Cigla; Rohan Thakker; Larry Matthies*

S4: Spotlights & Posters/Demos (1440-1540)

1440 Light Field Depth Estimation on Off-the-Shelf Mobile GPU, *Andre Ivan; Williem; In Kyu Park*

1442 Pseudo-Labels for Supervised Learning on Dynamic Vision Sensor Data, Applied to Object Detection Under Ego-Motion, *Nicholas F. Y. Chen*

1444 GPU Based Video Object Tracking on PTZ Cameras, *Cevahir Cigla; Kemal Emrecan Sahin; Fikret Alim*

1446 Analysis of Efficient CNN Design Techniques for Semantic Segmentation, *Alexandre Briot; Prashanth Viswanath; Senthil Yogamani*

1448 Design of a Reconfigurable 3D Pixel-Parallel Neuromorphic Architecture for Smart Image Sensor, *Pankaj Bhowmik; Md Jubaer Hossain Pantho; Marjan Asadinia; Christophe Bobda*

1450 Demo: PointAR: Augmented Reality Laser Pointer for Tele-Assistance, *Harald Haraldsson, Doron Tal, Karla Polo-Garcia, Serge Belongie*

1455 **Posters & Demo**

1540 Afternoon Break (Hall A; Halls 1-4)

S5: Oral 4 (1610-1720)

1610 **Invited Talk:** Hardware Design and Implementation of Convolutional Neural Networks, *Warren Gross (McGill Univ.)*

1700 KCNN: Extremely-Efficient Hardware Keypoint Detection With a Compact Convolutional Neural Network, *Paolo Di Febbo; Carlo Dal Mutto; Kinh Tieu; Stefano Mattoccia*

1720 **Awards & Closing Remarks**

Notes:

New Trends in Image Restoration and Enhancement

Organizers: Radu Timofte
Shuhang Gu
Luc Van Gool
Lei Zhang
Ming-Hsuan Yang
Cosmin Ancuti
Codruta O. Ancuti
Boaz Arad
Ohad Ben-Shahar
Jiqing Wu

Location: Room 257

Schedule: Full Day

0810 **Poster Setup** (all papers have poster panels for the whole day)

0830 **Opening Remarks**

0840 **Invited Talk:** High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs, *Ming-Yu Liu (NVIDIA)*

0910 **Invited Talk:** When Depth Estimation Meets Deep Learning, *Liang Lin (SenseTime / Sun Yat-sen Univ.)*

0940 WESPE: Weakly Supervised Photo Enhancer for Digital Cameras, *Andrey Ignatov; Nikolay Kobyshev; Radu Timofte; Kenneth Vanhoey; Luc Van Gool*

0950 Unsupervised Image Super-Resolution Using Cycle-in-Cycle Generative Adversarial Networks, *Yuan Yuan; Siyuan Liu; Jiawei Zhang; Yongbing Zhang; Chao Dong; Liang Lin*

1000 Morning Break (Hall A; Halls 1-4) **& Poster Session**

- DPW-SDNet: Dual Pixel-Wavelet Domain Deep CNNs for Soft Decoding of JPEG-Compressed Images, *Honggang Chen; Xiaohai He; Linbo Qing; Shuhua Xiong; Truong Q. Nguyen*
- Attribute Augmented Convolutional Neural Network for Face Hallucination, *Cheng-Han Lee; Kaipeng Zhang; Hu-Cheng Lee; Chia-Wen Cheng; Winston Hsu*
- Recursive Deep Residual Learning for Single Image Dehazing, *Yixin Du; Xin Li*
- Synthesized Texture Quality Assessment via Multi-Scale Spatial and Statistical Texture Attributes of Image and Gradient Magnitude Coefficients, *Alireza Golestaneh; Lina J. Karam*
- Learning Face Deblurring Fast and Wide, *Meiguang Jin; Michael Hirsch; Paolo Favaro*
- O-HAZE: A Dehazing Benchmark With Real Hazy and Haze-Free Outdoor Images, *Codruta O. Ancuti; Cosmin Ancuti; Radu Timofte; Christophe De Vleeschouwer*
- Large Receptive Field Networks for High-Scale Image Super-Resolution, *George Seif; Dimitrios Androutsos*
- Multi-Level Wavelet-CNN for Image Restoration, *Pengju Liu; Hongzhi Zhang; Kai Zhang; Liang Lin; Wangmeng Zuo*
- ComboGAN: Unrestrained Scalability for Image Domain Translation, *Asha Anoosheh; Eirikur Agustsson; Radu Timofte; Luc Van Gool*
- Image Super-Resolution via Progressive Cascading Residual Network, *Namhyuk Ahn; Byungkon Kang; Kyung-Ah Sohn*
- Deep Residual Network With Enhanced Upscaling Module for Super-Resolution, *Jun-Hyuk Kim; Jong-Seok Lee*
- Persistent Memory Residual Network for Single Image Super Resolution, *Rong Chen; Yanyun Qu; Kun Zeng; Jinkang Guo; Cuihua Li; Yuan Xie*
- Fully End-to-End Learning Based Conditional Boundary Equilibrium GAN With Receptive Field Sizes Enlarged for Single Ultra-High Resolution Image Dehazing, *Sehwan Ki; Hyeonjun Sim; Jae-Seok Choi; Saehun Kim; Munchurl Kim*
- Cycle-Dehaze: Enhanced CycleGAN for Single Image Dehazing, *Deniz Engin; Anıl Genç; Hazim Kemal Ekenel*
- IRGUN: Improved Residue Based Gradual Up-Scaling Network for Single Image Super Resolution, *Manoj Sharma; Rudrabha Mukhopadhyay; Avinash Upadhyay; Sriharsha Koundinya; Ankit Shukla; Santanu Chaudhury*
- 2D-3D CNN Based Architectures for Spectral Reconstruction From RGB Images, *Sriharsha Koundinya; Himanshu Sharma; Manoj Sharma; Avinash Upadhyay; Raunak Manekar; Rudrabha Mukhopadhyay; Abhijit Karmakar; Santanu Chaudhury*

1100 **Awards Ceremony**

1120 NTIRE 2018 Challenge on Single Image Super-Resolution: Methods and Results, *Radu Timofte; Shuhang Gu; Jiqing Wu; Luc Van Gool*

1130 A Fully Progressive Approach to Single-Image Super-Resolution, *Yifan Wang; Federico Perazzi; Brian*

McWilliams; Alexander Sorkine-Hornung; Olga Sorkine-Hornung; Christopher Schroers

1140 New Techniques for Preserving Global Structure and Denoising With Low Information Loss in Single-Image Super-Resolution, *Yijie Bei; Alexandru Damian; Shijia Hu; Sachit Menon; Nikhil Ravi; Cynthia Rudin*

1150 Efficient Module Based Single Image Super Resolution for Multiple Problems, *Dongwon Park; Kwanyoung Kim; Se Young Chun*

1200 Lunch (Hall A; Halls 1-4)

1330 **Invited Talk:** Copying and Editing Images, *William T. Freeman (MIT / Google)*

1400 NTIRE 2018 Challenge on Image Dehazing: Methods and Results, *Cosmin Ancuti; Codruta O. Ancuti; Radu Timofte*

1410 Multi-Scale Single Image Dehazing Using Perceptual Pyramid Deep Network, *He Zhang; Vishwanath Sindagi; Vishal M. Patel*

1420 High-Resolution Image Dehazing With Respect to Training Losses and Receptive Field Sizes, *Hyeonjun Sim; Sehwan Ki; Jae-Seok Choi; Soomin Seo; Saehun Kim; Munchurl Kim*

1430 Image Dehazing by Joint Estimation of Transmittance and Airlight Using Bi-Directional Consistency Loss Minimized FCN, *Ranjan Mondal; Sanchayan Santra; Bhabatosh Chanda*

1440 **Keynote Talk:** Computer Vision Technologies in City Brain, *Xian-Sheng Hua (Alibaba)*

1525 Afternoon Break (Hall A; Halls 1-4) **& Poster Session**

1630 NTIRE 2018 Challenge on Spectral Reconstruction From RGB Images, *Boaz Arad; Ohad Ben-Shahar; Radu Timofte*

1640 HSCNN+: Advanced CNN-Based Hyperspectral Recovery From RGB Images, *Zhan Shi; Chang Chen; Zhiwei Xiong; Dong Liu; Feng Wu*

1650 Reconstructing Spectral Images From RGB-Images Using a Convolutional Neural Network, *Tarek Stiebel; Simon Koppers; Philipp Seltsam; Dorit Merhof*

1700 **Invited Talk:** Metamer Sets, *Graham Finlayson (Univ. East Anglia, Spectral Edge Ltd)*

1730 **Closing Remarks**

Human Pose, Motion, Activities and Shape in 3D

Organizers: Grégory Rogez
Javier Romero

Location: Room 255 A

Schedule: Full Day

0850 **Welcome**

0900 **Invited Talk:** Pose or Attention for Human Activity Recognition? *Christian Wolf (INSA / INRIA)*

0930 **Invited Talk:** Monocular Motion and Performance Capture, *Christian Theobalt (MPII)*

1000 Morning Break (Hall A; Halls 1-4) **& Poster Session**

- Monocular RGB Hand Pose Inference From Unsupervised Refinable Nets, *Endri Dibra; Silvan Melchior; Ali Balkis; Thomas Wolf; Cengiz Oztireli; Markus Gross*
- Unsupervised Features for Facial Expression Intensity Estimation Over Time, *Maren Awiszus; Stella Graßhof; Felix Kuhnke; Jörn Ostermann*
- Deep Learning Whole Body Point Cloud Scans From a Single Depth Map, *Nolan Lunscher; John Zelek*
- HandyNet: A One-Stop Solution to Detect, Segment, Localize & Analyze Driver Hands, *Akshay Rangesh; Mohan M. Trivedi*

1100 **Invited Talk:** Analyzing Human Poses, Tracks, and Actions, *Deva Ramanan (Carnegie Mellon Univ.)*

1130 **Invited Talk:** TBA, *Yasser Sheik (Carnegie Mellon Univ.)*

1200 Lunch (Hall A; Halls 1-4)

1330 **Invited Talk:** TBA, *Michael J. Black (MPI-IS)*

1400 **Invited Talk:** TBA, *Kostas Daniilidis (Univ. of Pennsylvania)*

1430 **Invited Talk:** Inference of 3D Human Body Poses and Shapes, *Cordelia Schmid (INRIA / Google)*

1500 **Invited Talk:** TBA, *Francesc Moreno Noguer (IRI)*

1530 Afternoon Break (Hall A; Halls 1-4) **& Invited Posters**

1630 **Panel Discussion, Awards, & Closing Remarks**

Autonomous Driving

Organizers: Jose M. Alvarez
Fisher Yu
Ruigang Yang
Antonio M. Lopez
Andreas Geiger
Alan Yuille
Dinesh Manocha
Dequan Wang
David Vázquez
Hongdong Li
John Leonard
Jianxiong Xiao
Markus Enzweiler
Trevor Darrell
Tomas Pajdla

Location: Ballroom B & D

Schedule: Full Day

0855 **Welcome**

0900 **Invited Talk:** *Edwin Olson (Univ. of Michigan)*

0935 **Invited Talk:** *Will Maddern (Univ. of Oxford)*

1010 Morning Break (Hall A; Halls 1-4)

1045 **Invited Talk:** *Luc Vincent (Lyft)*

1120 **Invited Talk:** *Vladlen Koltun (Intel Labs)*

1200 Lunch (Hall A; Halls 1-4)

1310 **Invited Talk:** *Andreas Geiger (Univ. of Tübingen; MPI)*

1345 **Invited Talk:** *Kurt Keutzer (UC Berkeley)*

1420 **Invited Talk:** Secure Learning in Adversarial Environments, *Bo Li (Univ. of Illinois at Urbana-Champaign)*

1455 **Invited Talk:** *Andrej Karpathy (Tesla; Stanford Univ.)*

1530 Afternoon Break (Hall A; Halls 1-4) **& Poster Session**

- The ApolloScape Dataset for Autonomous Driving, *Xinyu Huang; Xinjing Cheng; Qichuan Geng; Binbin Cao; Dingfu Zhou; Peng Wang; Yuanqing Lin; Ruigang Yang*
- Scene Understanding Networks for Autonomous Driving Based on Around View Monitoring System, *Jeong Yeol Baek; Ioana Veronica Chelu; Livia Iordache; Vlad Paunescu; HyunJoo Ryu; Alexandru Ghiuta; Andrei Petreanu; Yun-Sung Soh; Andrei Leica; ByeongMoon Jeon*
- Training Deep Networks With Synthetic Data: Bridging the Reality Gap by Domain Randomization, *Jonathan Tremblay; Aayush Prakash; David Acuna; Mark Brophy; Varun Jampani; Cem Anil; Thang To; Eric Cameracci; Shaad Boochoon; Stan Birchfield*
- On the Iterative Refinement of Densely Connected Representation Levels for Semantic Segmentation, *Arantxa Casanova; Guillem Cucurull; Michal Drozdzal; Adriana Romero; Yoshua Bengio*
- Minimizing Supervision for Free-Space Segmentation, *Satoshi Tsutsui; Tommi Kerola; Shunta Saito; David J. Crandall*
- Error Correction for Dense Semantic Image Labeling, *Yu-Hui Huang; Xu Jia; Stamatios Georgoulis; Tinne Tuytelaars; Luc Van Gool*
- On the Importance of Stereo for Accurate Depth Estimation: An Efficient Semi-Supervised Deep Neural Network Approach, *Nikolai Smolyanskiy; Alexey Kamenev; Stan Birchfield*
- Accurate Deep Direct Geo-Localization From Ground Imagery and Phone-Grade GPS, *Shaohui Sun; Ramesh Sarukkai; Jack Kwok; Vinay Shet*
- Efficient and Safe Vehicle Navigation Based on Driver Behavior Classification, *Ernest Cheung; Aniket Bera; Dinesh Manocha*
- Detection of Distracted Driver Using Convolutional Neural Network, *Bhakti Baheti; Suhas Gajre; Sanjay Talbar*
- Classifying Group Emotions for Socially-Aware Autonomous Vehicle Navigation, *Aniket Bera; Tanmay Randhavane; Austin Wang; Dinesh Manocha; Emily Kubin; Kurt Gray*
- AutonoVi-Sim: Autonomous Vehicle Simulation Platform With Weather, Sensing, and Traffic Control, *Andrew Best; Sahil Narang; Lucas Pasqualin; Daniel Barber; Dinesh Manocha*
- Learning Hierarchical Models for Class-Specific Reconstruction From Natural Data, *Arun CS Kumar; Suchendra M. Bhandarkar; Mukta Prasad*
- Subset Replay Based Continual Learning for Scalable Improvement of Autonomous Systems, *Pratik Prabhanjan Brahma; Adrienne Othon*

1630 **Panel**, *Moderator: Jose Alvarez*

1730 **Closing Remarks**

2030 **Reception**

Medical Computer Vision and Health Informatics

Organizers: Le Lu
Leo Grady
Tal Arbel
Xiaosong Wang
Nicolas Padoy

Location: Room 355 D

Schedule: Full Day

0800 **Invited Talk:** Quantitative Medical Image Analysis and Image-Based Computational Physiology, *Alex Frangi (Univ. of Sheffield)*

0830 **Invited Talk:** Regression Modeling on Manifolds for Analyzing Brain Shape and Functional Networks, *Tom Fletcher (Univ. of Utah)*

0900 **Invited Talk:** Vision for Healthcare (Dementia), *Francois Brémond (INRIA)*

0930 **Invited Talk:** Towards Ambient Intelligence in AI-Assisted Hospitals, *Serena Yeung (Stanford Univ.)*

1000 Morning Break (Hall A; Halls 1-4)

1015 **Invited Talk:** Efficient Statistical Shape Analysis Based on Geometric Deformations, *MiaoMiao Zhang (Lehigh Univ.)*

1045 **Invited Talk:** Deep Learning for Biomedical Imaging: Can We Get Better, Higher or Faster? *Tammy Riklin-Raviv (Ben Gurion Univ.)*

1115 **Invited Talk:** Semi-Supervised Learning for Recognition in Endoscopic Videos, *Nicolas Padoy (Univ. of Strasbourg)*

1145 **Invited Talk:** Quantifying Surgery From Video Data, *Greg Hager (Johns Hopkins Univ.)*

Lunch (Hall A; Halls 1-4)

1330 **Invited Talk:** Using Computer Vision for Computer-Aided Surgery by Augmenting Laparoscopy With Preoperative Image Data, *Adrien Bartoli (Université Clermont Auvergne)*

1400 **Invited Talk:** Beyond Naive Labeling: Active Learning and Passive Annotations for Medical Imaging, *Raphael Sznitman (Univ. of Bern)*

1430 **Invited Talk:** Blood Cell Reconstruction, Detection, Classification and Counting in Holographic Images, *Rene Vidal (Johns Hopkins Univ.)*

1500 **Invited Talk:** Mine Deeper & Learn Wider: A Perspective on Distilling Radiological Reports for Chest X-Ray Analysis, *Xiaosong Wang (NIH)*

1530 Afternoon Break (Hall A; Halls 1-4)

1545 **Invited Talk:** Personalized Blood Flow Simulation From an Image-Derived Model: Changing the Paradigm for Cardiovascular Diagnostics, *Leo Grady (HeartFlow)*

1615 **Invited Talk:** Bi-Modal and Semi-Supervised Models for Deep Learning in Non-Curated Medical Image Collections, *Mehdi Moradi (IBM Almaden Research Center)*

1645 **Invited Talk:** Lesion Detection and Localization in Mammography via Multi-Task Learning, *Nicolas Thome (Conservatoire National des Arts et Métiers)*

1715 **Invited Talk:** Computational Pathology at Scale: Changing Clinical Practice One Petabyte at a Time, *Thomas Fuchs (Memorial Sloan Kettering Cancer Center)*

Language and Vision

Organizers: Andrei Barbu
Tao Mei
Siddharth Narayanaswamy
Jiebo Luo
Dan Gutfreund
Yevgeni Berzak

Location: Room 255 B

Schedule: Full Day

0920 **Welcome**

0930 **Invited Talk:** *Anthony Cohn (Univ. of Leeds)*

1015 **Invited Talk:** *Roger Levy (MIT)*

1100 Morning Break (Hall A; Halls 1-4)

1115 **Flash Presentations**

1200 Lunch (Hall A; Halls 1-4) **& Poster Session**

1430 **Invited Talk:** *Kevin Murphy (Google)*

1515 **Invited Talk:** *Kristen Graumann (Univ. of Texas at Austin)*

1600 Afternoon Break (Hall A; Halls 1-4)

1615 **Invited Talk:** *Mark Yatskar (Univ. of Washington)*

1700 **Invited Talk:** TBA

1745 **Closing Remarks**

1830 **Dinner**

Brave New Ideas for Video Understanding

Organizers: Basura Fernando
Efstratios Gavves
Rahul Sukthankar
Lorenzo Torresani
Jan van Gemert
Andrew Zisserman

Location: Room 251 D-F

Schedule: Full Day

0900 **Welcome**

0910 **Invited Talk:** *Iasonas Kokkinos (Facebook, UCL)*

0940 **Invited Talk:** *Honglak Lee (U. Michigan, Google Brain)*

1010 Morning Break (Hall A; Halls 1-4)

1050 Temporal Reasoning in Videos Using Convolutional Gated Recurrent Units, *Debidatta Dwibedi; Pierre Sermanet; Jonathan Tompson*

1110 Temporal 3D ConvNets Using Temporal Transition Layer, *Ali Diba; Mohsen Fayyaz; Vivek Sharma; A. Hossein Karami; M. Mahdi Arzani; Rahman Yousefzadeh; Luc Van Gool*

1130 **Invited Talk:** *TBA*

1200 Lunch (Hall A; Halls 1-4)

1430 **Invited Talk:** *Cees Snoek (Univ. of Amsterdam)*

1500 ContextVP: Fully Context-Aware Video Prediction, *Wonmin Byeon; Qin Wang; Rupesh Kumar Srivastava; Petros Koumoutsakos*

1520 Afternoon Break (Hall A; Halls 1-4) **& Poster Session**

- Towards an Unequivocal Representation of Actions, *Michael Wray; Davide Moltisanti; Dima Damen*
- Unsupervised Deep Representations for Learning Audience Facial Behaviors, *Suman Saha; Rajitha Navarathna; Leonhard Helminger; Romann M. Weber*
- I Have Seen Enough: A Teacher Student Network for Video Classification Using Fewer Frames, *Shweta Bhardwaj; Mitesh M. Khapra*

1630 **Invited Talk:** *Ivan Laptev (INRIA)*

1700 **Closing Remarks**

Deep-Vision

Organizers: Jose M. Alvarez
Nathan Silberman
Dhruv Batra
Yann LeCun

Location: Room 254

Schedule: Full Day

0915 **Welcome**

0920 **Invited Talk:** *Kevin Murphy (Google)*

1000 Morning Break (Hall A; Halls 1-4)

1030 **Invited Talk:** *Josef Civic (INRIA)*

1105 **Invited Talk:** *Adriana Romero (Facebook AI)*

1140 **Invited Talk:** *Olga Russakovsky (Princeton)*

1225 Lunch (Hall A; Halls 1-4)

1400 **Invited Talk:** *Vittorio Ferrari (Google)*

1435 **Invited Talk:** *Chris Re (Stanford)*

1510 **Invited Talk:** *Devi Parik (Georgia Tech and Facebook AI)*

1545 Afternoon Break (Hall A; Halls 1-4) **& Poster Session**

Vision Meets Cognition: Functionality, Physics, Intentionality and Causality

Organizers: Yixin Zhu
Yibiao Zhao
Jiajun Wu
Lap-Fai Yu
Ping Wei
Chenfanfu Jiang
Tao Gao

Location: Room 155 B

Schedule: Full Day

0950 **Welcome**

1000 **Invited Talk:** *Igor Mordatch (OpenAI)*

1040 **Invited Talk:** *Patrick Shafto (Rutgers Univ. - Newark)*

1120 **Invited Talk:** *Tomer Ullman (Harvard Univ.)*

1200 Lunch (Hall A; Halls 1-4)

1340 Spotlight Presentation

1410 **Invited Talk:** *Ashutosh Saxena (Stanford Univ.)*

1450 **Invited Talk:** *William T. Freeman (MIT)*

1530 Afternoon Break (Hall A; Halls 1-4) **& Poster Session**
1600 **Invited Talk:** *Caiming Xiong (Salesforce Research)*
1640 **Invited Talk:** *Dileep George (Vicarious AI)*
1720 **Closing Remarks**
1730 **Private Party**

Perception Beyond the Visible Spectrum

Organizers: Riad I. Hammoud
Michael Teutsch
Yi Ding

Location: Room 150 A-C

Schedule: Full Day

0830 **Welcome & Opening Remarks**
0845 **Keynote Talk:** Generative Adversarial Networks, *Ian Goodfellow (Google Brain)*
0930 **Keynote Talk:** RGB+: Using Near-Infrared (NIR) to improve Computational Photography Applications, *Sabine Suesstrunk (EPFL)*

1015 Morning Break (Hall A; Halls 1-4)

S1: Advances in Generative Adversarial Networks and Image Enhancement (1045-1200)

1045 Generating Visible Spectrum Images From Thermal Infrared, *Amanda Berg; Jörgen Ahlberg; Michael Felsberg*
1100 IR2VI: Enhanced Night Environmental Perception by Unsupervised Thermal Image Translation, *Shuo Liu; Vijay John; Erik Blasch; Zheng Liu; Ying Huang*
1115 Path Orthogonal Matching Pursuit for Sparse Reconstruction and Denoising of SWIR Maritime Imagery, *Timothy Doster; Tegan Emerson; Colin Olson*
1130 Deep Learning Based Single Image Dehazing, *Patricia L. Suárez; Angel D. Sappa; Boris X. Vintimilla; Riad I. Hammoud*
1145 Generative Adversarial Networks for Depth Map Estimation From RGB Video, *Kin Gwn Lore; Kishore Reddy; Michael Giering; Edgar A. Bernal*

1200 Lunch (Hall A; Halls 1-4)

1330 **Keynote Talk:** Hyperspectral Imaging Remote Sensing: Progress and Challenges, *Dimitris Manolakis (MIT Lincoln Lab)*

S2: Advances in Detection/Tracking and Recognition in Infrared (1415-1545)

1415 On the Impact of Parallax Free Colour and Infrared Image Co-Registration to Fused Illumination Invariant Adaptive Background Modelling, *Michael Loveday; Toby P. Breckon*
1430 Integrated Learning and Feature Selection for Deep Neural Networks in Multispectral Images, *Anthony Ortiz; Alonso Granados; Olac Fuentes; Christopher Kiekintveld; Dalton Rosario; Zachary Bell*
1445 A Comprehensive Solution for Deep-Learning Based Cargo Inspection to Discriminate Goods in Containers, *Jiahang Che; Yuxiang Xing; Li Zhang*
1500 Cross-Domain Hallucination Network for Fine-Grained Object Recognition, *Jin-Fu Lin; Yen-Liang Lin; Erh-Kan King; Hung-Ting Su; Winston H. Hsu*
1515 Deep Convolutional Neural Networks With Integrated Quadratic Correlation Filters for Automatic Target Recognition, *Brian Millikan; Hassan Foroosh; Qiyu Sun*
1530 An Online and Flexible Multi-Object Tracking Framework Using Long Short-Term Memory, *Xingyu Wan; Jinjun Wang; Sanping Zhou*

1545 Afternoon Break (Hall A; Halls 1-4)

S3: Synthetic-Aperture-Radar and Long-Wave Infrared Data Exploitation (1615-1645)

1615 Polarimetric Synthetic-Aperture-Radar Change-Type Classification With a Hyperparameter-Free Open-Set Classifier, *Mark W. Koch; R. Derek West; Robert Riley; Tu-Thach Quach*
1630 POL-LWIR Vehicle Detection: Convolutional Neural Networks Meet Polarised Infrared Sensors, *Marcel Sheeny; Andrew Wallace; Mehryar Emambakhsh; Sen Wang; Barry Connor*
1645 **Closing Remarks**

Learnt Image Compression

Organizers: George Toderici
Radu Timofte
William T. Freeman
Michele Covell
Balu Adsumili
Wenzhe Shi
Lucas Theis
Johannes Ballé
Eirikur Agustsson
Nick Johnston

Location: Room 250 C

Schedule: Half Day - Morning

0810 **Poster Setup** (all papers have poster panels for the whole day)
0830 **Opening Remarks**
0845 **Invited Talk:** *Jim Bankoski (Google)*
0905 **Invited Talk:** *Jens Ohm (RWTH Aachen)*
0925 **Invited Talk:** *Oren Rippel (WaveOne)*
0945 **Invited Talk:** *Ebrahimi Touradj (EPFL)*

1005 Morning Break (Hall A; Halls 1-4) **& Poster Session**

1105 **Challenge Presentation —** Fastest
1115 **Challenge Presentation —** Third Place
1125 **Challenge Presentation —** Second Place
1135 **Challenge Presentation —** First Place
1145 **Panel Discussion**
1225 **Awards Ceremony**

1235 Lunch (Hall A; Halls 1-4) **& Discussions**

Fine-Grained Instructional Video Understanding

Organizers: Jason Corso
Ivan Laptev
Josef Sivic
Luowei Zhou

Location: Room 151 G

Schedule: Half Day - Afternoon

1330 **Welcome**
1340 **Invited Talk:** Finding "It": Weakly Supervised Reference-Aware Visual Grounding in Instructional Video, *Animesh Garg (Stanford Univ.)*
1405 **Invited Talk:** Leveraging Motoric Information for Fine-Grained Action Understanding, *Cornelia Fermuller (Univ. of Maryland, College Park)*
1430 **Invited Talk:** Learning Grounded Game Play, *Jeffrey Siskind (Purdue Univ.)*

1455 Afternoon Break (Hall A; Halls 1-4) **& Poster Session**

1600 **Invited Talk:** RoboBrain: Learning Actionable Representations from Videos, *Ashutosh Saxena (Stanford Univ. / Caspar.AI)*
1625 **Invited Talk:** TBA, *Abhinav Gupta (Carnegie Mellon Univ.)*
1650 **Panel:** Future of Instructional Video in Computer Vision

Large-Scale Landmark Recognition: A Challenge

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

Location: Room 251 B-C

Schedule: Half Day - Afternoon

1340 **Welcome**
1350 **Invited Talk:** *Josef Sivic (INRIA)*
1430 **Challenge Overview**

S1: Recognition Challenge (1450-1530)

1450 **Challenge Presentation —** Recognition Challenge Winner
1510 **Challenge Presentation —** Recognition Challenge Participant(s)

1530 Break (Hall A; Halls 1-4)

S2: Retrieval Challenge (1600-1640)

1600 **Challenge Presentation —** Retrieval Challenge Winner
1620 **Challenge Presentation —** Retrieval Challenge Participant(s)
1640 **Invited Talk:** *Herve Jegou (Facebook AI Research)*
1720 **Closing Remarks**

DAVIS Challenge on Video Object Segmentation

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

Location: Room 255 D

Schedule: Half Day - Afternoon

1330 **Welcome**
1340 **Invited Talk:** *Fuxin Li (Oregon State Univ.)*
1400 **Challenge Presentation —** Fourth Place, *4th Place Challenge*
1410 **Challenge Presentation —** Third Place
1420 **Invited Talk:** *Iasonas Kokkinos (Facebook AI Research, UCL)*
1440 **Challenge Presentation —** Second Place
1455 **Challenge Presentation —** First Place
1510 **Invited Talk:** *Joon-Young Lee (Adobe Research)*
1530 Afternoon Break (Hall A; Halls 1-4) **& Posters**
1645 **Invited Talk:** *TBA (Google AI Perception)*
1705 **Awards & Closing Remarks**

Bridging the Gap Between Computational Photography and Visual Recognition: The UG^2 Prize Challenge

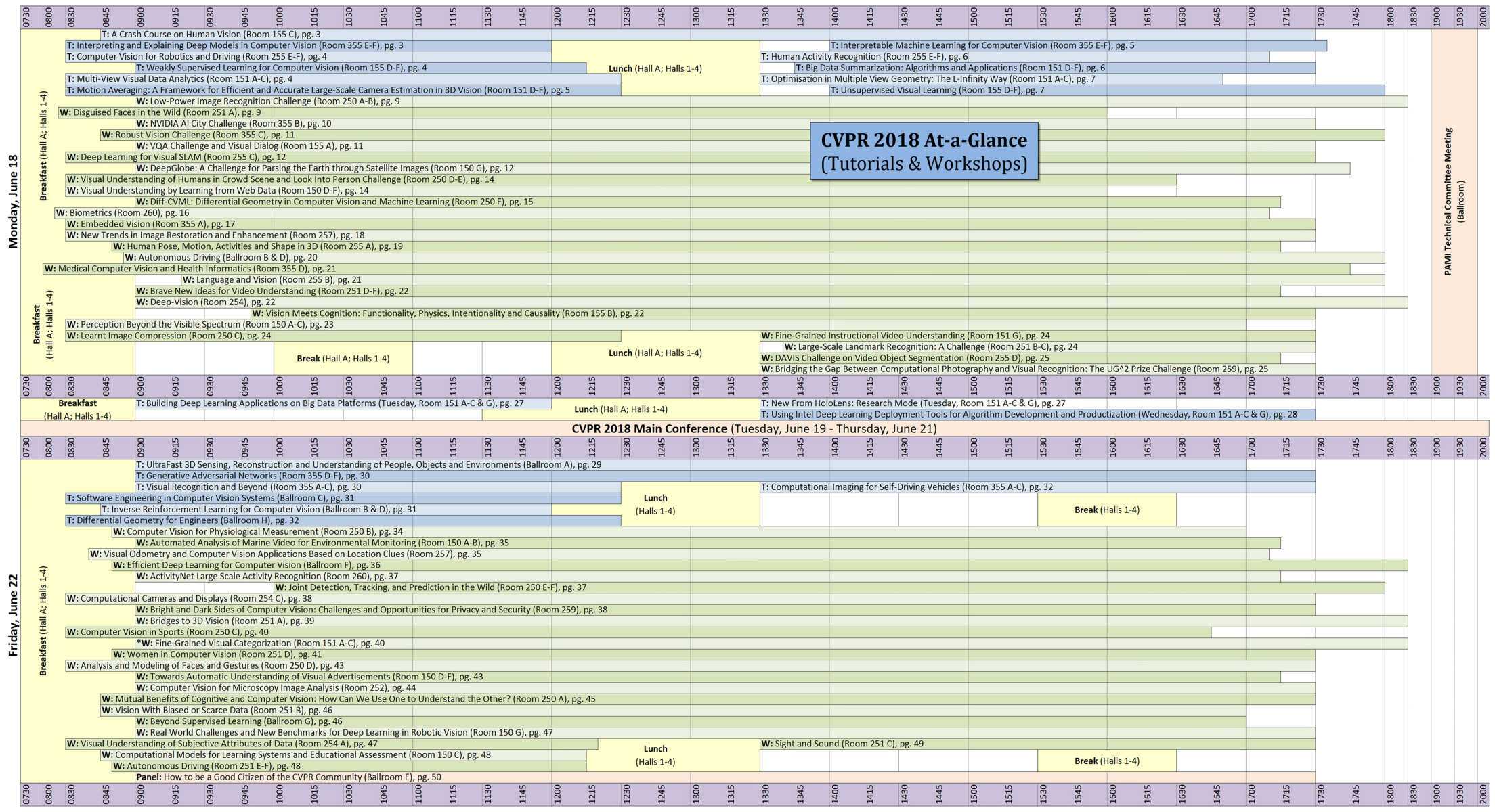
Organizers: Walter J. Scheirer
Rosaura Vidal Mata
Sreya Banerjee

Location: Room 259

Schedule: Half Day - Afternoon

1330 **Keynote Talk:** Prize Challenges and IARPA's Priorities, *Christopher Boehnen (IARPA)*
1400 **Overview of the UG^2 Challenge**
1430 **Keynote Talk:** Enhancement of Overhead and Aerial Imagery Using 3D and Deep Learning, *Anthony Hoogs (Kitware)*
1500 **Keynote Talk:** Computer Vision for Autonomous Aerial Vehicles, *Camillo J. Taylor (Univ. of Pennsylvania)*
1530 Afternoon Break (Hall A; Halls 1-4)
1600 **Announcement of Winners**
1615 **Challenge 1 Presentation — Runner Up:** Image Enhancement to Facilitate Manual Inspection
1630 **Challenge 1 Presentation — Winner:** Image Enhancement to Facilitate Manual Inspection
1645 **Challenge 2 Presentation — Runner Up:** Image Enhancement to Improve Automatic Object Recognition
1700 **Challenge 2 Presentation — Winner:** Image Enhancement to Improve Automatic Object Recognition
1715 **Closing Remarks**

Notes:


CVPR 2018 At-a-Glance
(Tutorials & Workshops)
Monday, June 18
Breakfast (Hall A; Halls 1-4)
T: A Crash Course on Human Vision (Room 155 C), pg. 3
T: Interpreting and Explaining Deep Models in Computer Vision (Room 355 E-F), pg. 3
T: Computer Vision for Robotics and Driving (Room 255 E-F), pg. 4
T: Weakly Supervised Learning for Computer Vision (Room 155 D-F), pg. 4
T: Multi-View Visual Data Analytics (Room 151 A-C), pg. 4
T: Motion Averaging: A Framework for Efficient and Accurate Large-Scale Camera Estimation in 3D Vision (Room 151 D-F), pg. 5
W: Low-Power Image Recognition Challenge (Room 250 A-B), pg. 9
W: Disguised Faces in the Wild (Room 251 A), pg. 9
W: NVIDIA AI City Challenge (Room 355 B), pg. 10
W: Robust Vision Challenge (Room 355 C), pg. 11
W: VQA Challenge and Visual Dialog (Room 155 A), pg. 11
W: Deep Learning for Visual SLAM (Room 255 C), pg. 12
W: DeepGlobe: A Challenge for Parsing the Earth through Satellite Images (Room 150 G), pg. 12
W: Visual Understanding of Humans in Crowd Scene and Look Into Person Challenge (Room 250 D-E), pg. 14
W: Visual Understanding by Learning from Web Data (Room 150 D-F), pg. 14
W: Diff-CVML: Differential Geometry in Computer Vision and Machine Learning (Room 250 F), pg. 15
W: Biometrics (Room 260), pg. 16
W: Embedded Vision (Room 355 A), pg. 17
W: New Trends in Image Restoration and Enhancement (Room 257), pg. 18
W: Human Pose, Motion, Activities and Shape in 3D (Room 255 A), pg. 19
W: Autonomous Driving (Ballroom B & D), pg. 20
W: Medical Computer Vision and Health Informatics (Room 355 D), pg. 21
W: Language and Vision (Room 255 B), pg. 21
W: Brave New Ideas for Video Understanding (Room 251 D-F), pg. 22
W: Deep-Vision (Room 254), pg. 22
W: Vision Meets Cognition: Functionality, Physics, Intentionality and Causality (Room 155 B), pg. 22
W: Perception Beyond the Visible Spectrum (Room 150 A-C), pg. 23
W: Learnt Image Compression (Room 250 C), pg. 24
Lunch (Hall A; Halls 1-4)
T: Interpretable Machine Learning for Computer Vision (Room 355 E-F), pg. 5
T: Human Activity Recognition (Room 255 E-F), pg. 6
T: Big Data Summarization: Algorithms and Applications (Room 151 D-F), pg. 6
T: Optimisation in Multiple View Geometry: The L-Infinity Way (Room 151 A-C), pg. 7
T: Unsupervised Visual Learning (Room 155 D-F), pg. 7
PAMI Technical Committee Meeting (Ballroom)
Breakfast (Hall A; Halls 1-4)
Break (Hall A; Halls 1-4)
Lunch (Hall A; Halls 1-4)
W: Fine-Grained Instructional Video Understanding (Room 151 G), pg. 24
W: Large-Scale Landmark Recognition: A Challenge (Room 251 B-C), pg. 24
W: DAVIS Challenge on Video Object Segmentation (Room 255 D), pg. 25
W: Bridging the Gap Between Computational Photography and Visual Recognition: The UG^2 Prize Challenge (Room 259), pg. 25
Breakfast (Hall A; Halls 1-4)
T: Building Deep Learning Applications on Big Data Platforms (Tuesday, Room 151 A-C & G), pg. 27
Lunch (Hall A; Halls 1-4)
T: New From HoloLens: Research Mode (Tuesday, Room 151 A-C & G), pg. 27
T: Using Intel Deep Learning Deployment Tools for Algorithm Development and Productization (Wednesday, Room 151 A-C & G), pg. 28
CVPR 2018 Main Conference (Tuesday, June 19 - Thursday, June 21)
Friday, June 22
Breakfast (Hall A; Halls 1-4)
T: UltraFast 3D Sensing, Reconstruction and Understanding of People, Objects and Environments (Ballroom A), pg. 29
T: Generative Adversarial Networks (Room 355 D-F), pg. 30
T: Visual Recognition and Beyond (Room 355 A-C), pg. 30
T: Software Engineering in Computer Vision Systems (Ballroom C), pg. 31
T: Inverse Reinforcement Learning for Computer Vision (Ballroom B & D), pg. 31
T: Differential Geometry for Engineers (Ballroom H), pg. 32
W: Computer Vision for Physiological Measurement (Room 250 B), pg. 34
W: Automated Analysis of Marine Video for Environmental Monitoring (Room 150 A-B), pg. 35
W: Visual Odometry and Computer Vision Applications Based on Location Clues (Room 257), pg. 35
W: Efficient Deep Learning for Computer Vision (Ballroom F), pg. 36
W: ActivityNet Large Scale Activity Recognition (Room 260), pg. 37
W: Joint Detection, Tracking, and Prediction in the Wild (Room 250 E-F), pg. 37
W: Computational Cameras and Displays (Room 254 C), pg. 38
W: Bright and Dark Sides of Computer Vision: Challenges and Opportunities for Privacy and Security (Room 259), pg. 38
W: Bridges to 3D Vision (Room 251 A), pg. 39
W: Computer Vision in Sports (Room 250 C), pg. 40
*W: Fine-Grained Visual Categorization (Room 151 A-C), pg. 40
W: Women in Computer Vision (Room 251 D), pg. 41
W: Analysis and Modeling of Faces and Gestures (Room 250 D), pg. 43
W: Towards Automatic Understanding of Visual Advertisements (Room 150 D-F), pg. 43
W: Computer Vision for Microscopy Image Analysis (Room 252), pg. 44
W: Mutual Benefits of Cognitive and Computer Vision: How Can We Use One to Understand the Other? (Room 250 A), pg. 45
W: Vision With Biased or Scarce Data (Room 251 B), pg. 46
W: Beyond Supervised Learning (Ballroom G), pg. 46
W: Real World Challenges and New Benchmarks for Deep Learning in Robotic Vision (Room 150 G), pg. 47
W: Visual Understanding of Subjective Attributes of Data (Room 254 A), pg. 47
W: Computational Models for Learning Systems and Educational Assessment (Room 150 C), pg. 48
W: Autonomous Driving (Room 251 E-F), pg. 48
Panel: How to be a Good Citizen of the CVPR Community (Ballroom E), pg. 50
Lunch (Halls 1-4)
T: Computational Imaging for Self-Driving Vehicles (Room 355 A-C), pg. 32
Break (Halls 1-4)
W: Sight and Sound (Room 251 C), pg. 49
Lunch (Halls 1-4)
Break (Halls 1-4)

Tuesday, June 19

0730–1830 Registration (South Lobby)

0730–0900 Breakfast (Hall A; Halls 1-4)

1000–1045 Morning Break (Hall A; Halls 1-4)

1130–1330 Lunch (Hall A; Halls 1-4)

1430–1500 Afternoon Break (Hall A; Halls 1-4)

Building Deep Learning Applications on Big Data Platforms

Organizers: Jason (Jinquan) Dai

Location: Room 151 A-C & G

Time: 0900-1200 (Half Day — Morning)

Description: Recent breakthroughs in artificial intelligence applications have brought deep learning to the forefront of new generations of data analytics. In this tutorial, we will present the practice and design tradeoffs for building large-scale deep learning applications (such as computer vision and NLP) for production data and workflow on Big Data platforms. In particular, we will provide an overview of emerging deep learning frameworks for Big Data (e.g., BigDL, TensorFlow-on-Spark, Deep Learning Pipelines for Spark, etc.), present the underlying distributed systems and algorithms, and discuss innovative data analytics + AI application pipelines (with a focus on computer vision models and use cases) for Big Data platforms and workflows.

Schedule:

0900 Motivation
0910 Overview
0930 Analytics Zoo for Spark and BigDL

1000 Morning Break

1030 Distributed Training and Inference
1100 Advanced Applications
1130 Real-World Applications
1150 Q&A

New From HoloLens: Research Mode

Organizers: Marc Pollefeys
Georg Klein
Shivkumar Swaminathan
Larry Wall
Sudipta Sinha
Johannes Schönberger

Location: Room 151 A-C & G

Time: 1330-1450 (Half Day — Afternoon)

Description: The HoloLens is the world's leading augmented reality headset, but it's also a potent computer vision research device. Application code can access audio and video streams and surface meshes, all in a world coordinate space maintained by HoloLens' highly accurate head-tracking. This shorttutorial will dive into the just-released "Research Mode" capability of HoloLens. We will show you how to access therawhead-tracking and depth sensor data streams, andmake use of the intrinsics and extrinsicsof each stream.We will also be demonstrating recent advances intime of flight depth-sensing technologies.

Schedule:

1330 Introduction to HoloLens
1345 Research Mode: Gettingyour Hands on the Sensor Streams
1415 Research Mode in Use–Demos and Videos
1445 Sneak Peekatrecent Advances in Time-of-Flight Depth Sensing

Notes:

Wednesday, June 20

0730–1830 Registration (South Lobby)

0730–0900 Breakfast (Hall A; Halls 1-4)

1000–1045 Morning Break (Hall A; Halls 1-4)

1130–1330 Lunch (Hall A; Halls 1-4)

1430–1500 Afternoon Break (Hall A; Halls 1-4)

Using Intel Deep Learning Deployment Tools for Algorithm Development and Productization

Organizers: Alexander Bovyrin
Yury Gorbachev
Vadim Pisarevsky

Location: Room 151 A-C & G

Time: 1330-1730 (Half Day — Afternoon)

Description: Deep Learning based algorithms are very resource demanding and compute intensive tasks. Selection of deployment tools is important choice that has to be made whenever algorithm is ready for production. In practice, selection of tools not only impacts speed, but also quality of the final solution. Hence, knowledge about target platform and features that are supported by enabling tools at algorithm selection stage can be critical.

Schedule:

1330 Architecture of Intel Deep Learning Deployment Tools
1415 Cross-Platform Portability for DL Deployment
1445 Most Important Features (Topology Specific Optimizations, Dynamic Batching, Support for Various Levels of Precision, Inference Engine Integration Into OpenCV)
1530 Afternoon Break
1600 Examples of Deployment Tools Usage
1620 CV SDK Models Zoo
1640 Planned CV SDK Features
1700 Overview of R&D in Deep Nets Optimization

Notes:

Friday, June 22

0730–1830 Registration (North Lobby)

0730–0900 Breakfast (Halls 1-4)

1000–1100 Morning Break (Halls 1-4)

1200–1330 Lunch (Halls 1-4)

1530–1630 Afternoon Break (Halls 1-4)

UltraFast 3D Sensing, Reconstruction and Understanding of People, Objects and Environments

Organizers: Sean Fanello, Julien Valentin, Jonathan Taylor, Christoph Rhemann, Adarsh Kowdle, Jürgen Sturm, Christine Kaeser-Chen, Pavel Pidlypenskyi, Rohit Pandey, Sameh Khamis, David Kim, Mingsong Dou, Kaiwen Guo, Danhang Tang, Shahram Izadi

Location: Ballroom A

Time: 0900-1700 (Full Day)

Description: The emergence of virtual and augmented reality has increased the demand of robust systems for 3D capture, reconstruction and understanding. Designing such systems involves developing high quality sensors and efficient algorithms that can leverage new and existing technologies. To this end, we design depth sensors with two properties that drastically simplify the problem of fusing incomplete sensor data. First, an UltraFast (~1000 fps) depth stream that significantly reduces the frame to frame motion. Second, by allowing multiple sensors to be easily combined (without interference) we can eliminate occlusions. As a result, we have developed a wide range of efficient algorithms for reconstruction, tracking and understanding that are designed to work with this technology. In this tutorial we will walk the audience through the ins and outs of building such a stack from the ground up. From building such a sensor to applications for mixed reality with particular emphasis on high speed 3D capture systems.

Schedule:

0900 Introduction
0910 Depth Sensors and Algorithms: What, When, Where, *Adarsh Kowdle*
0940 Triangulation Methods: From Theory to Mobile, *Christoph Rhemann*
1010 UltraFast Machine Learning for Triangulation Methods, *Sean Fanello*

1030 Morning Break

1100 Deep Learning to Break 1/30th Subpixel Precision, *Sameh Khamis*
1115 Active Stereo Net: Self-Supervised End-to-End Training for Active Stereo Systems, *Yinda Zhang*
1130 Non-Linear Optimization Methods, *Jonathan Taylor*

1200 Lunch

1300 Localization and Mapping - ARCore, *Konstantine Tsotsos, Zack Moratto*
1345 Machine Learning for Camera Relocalization, *Julien Valentin*
1400 3D Scene Understanding, *Juergen Sturm*
1415 Semantic Labeling, *Thomas Funkhouser*
1430 Parametric Tracking, *Anastasia Tkach*
1450 Non-Rigid Reconstruction of Humans, *Mingsong Dou, Kaiwen Guo*

1530 Afternoon Break

1600 LookinGood - Free Viewpoint Volumetric Videos, *Rohit Pandey*
1615 Fast Machine Learning on Embedded Devices, *Christine Kaeser-Chen*
1630 Digital AR, *David Kim*

Generative Adversarial Networks

Organizers: Jun-Yan Zhu
Taesung Park
Mihaela Rosca
Phillip Isola
Ian Goodfellow.

Location: Room 355 D-F

Time: 0900-1730 (Full Day)

Description: Generative adversarial networks (GANs) have been at the forefront of research on generative models in the last couple of years. GANs have been used for image generation, image processing, image synthesis from captions, image editing, visual domain adaptation, data generation for visual recognition, and many other applications, often leading to state of the art results. This tutorial aims to provide a broad overview of generative adversarial networks, mainly including the following three parts: (1) theoretical foundations such as basic concepts, mechanisms, and theoretical considerations, (2) best practices of the current state-of-the-art GAN and conditional GAN models, including network architectures, objective functions, and other training tricks, and (3) computer vision applications including visual domain adaptation, image processing (e.g., restoration, inpainting, super-resolution), image synthesis and manipulation, video prediction and generation, 3D modeling, synthetic data generation for visual recognition, robotic learning, and so on.

Schedule:

0900 Introduction to Generative Adversarial Networks, *Ian Goodfellow (Google Brain)*

0930 Paired Image-to-Image Translation, *Phillip Isola (MIT)*

1000 Unpaired Image-to-Image Translation, *Taesung Park (UC Berkeley and Jun-Yan Zhu, MIT)*

1030 Morning Break & Live Demo

1100 Can GANs Actually Learn the Distribution? Some Obstacles, *Sanjeev Arora (Princeton)*

1145 TBA, *Emily Denton (NYU)*

1215 Lunch

1330 Autoencoder, VAE, and GANs, *Mihaela Rosca (Deepmind)*

1400 TBA, *Ming-Yu Liu (NVIDIA)*

1430 Adversarial Domain Adaptation, *Judy Hoffman (UC Berkeley)*

1500 Afternoon Break & Live Demo

1530 TBA, *Abhinav Gupta (CMU)*

1600 Generative Adversarial Imitation Learning, *Stefano Ermon (Stanford)*

1630 Video Generation and Prediction, *Carl Vondrick (Columbia Univ.)*

1700 TBA, *Alexei A. Efros (UC Berkeley)*

Visual Recognition and Beyond

Organizers: Kaiming He
Ross Girshick
Alex Kirillov
Georgia Gkioxari
Justin Johnson

Location: Room 355 A-C

Time: 0900-1230 (Half Day — Morning)

Description: This tutorial covers topics at the frontier of research on visual recognition and its use in higher-level tasks (the beyond). The first three talks cover methods and principles behind image classification, object detection, instance segmentation, and semantic segmentation. The final two talks explore newly emerging tasks involving the use of visual recognition for action and reasoning.

Schedule:

0900 Learning Deep Representations for Visual Recognition, *Kaiming He*

0930 The Generalized R-CNN Framework for Object Detection, *Ross Girshick*

1000 Panoptic Segmentation: Unifying Semantic and Instance Segmentations, *Alex Kirillov*

1030 Morning Break

1130 Embodied Vision, *Georgia Gkioxari*

1200 TBA, *Justin Johnson*

Software Engineering in Computer Vision Systems

Organizers: David Doria

Location: Ballroom C

Time: 0830-1230 (Half Day — Morning)

Description: Each year top computer vision researchers from around the world gather at CVPR to present and discuss recent developments and results from both academic and industrial efforts. Topics include things such as new procedures for training neural networks, evaluation results on standard classification datasets, and demonstrations of applications enabled by the new algorithms. While the algorithms themselves are very exciting and critical components, when handed a new algorithm there is a non-trivial amount of work required to design, develop, run, and maintain a full software system around the algorithm. In fact, many teams in industry have software engineers, not computer vision experts, as many of their members. This workshop is intended to shine a light on this typically unpublicized part of the process with hopes to share best practices, expose common hurdles, and explain the complexity to computer vision researchers who may not be familiar with large system development.

Schedule:

0830 Accelerating Algorithm Development, Evaluation, and Deployment by Providing Frameworks, *David Doria (HERE)*

0900 Reproducibility – An Industrial Practice, *Jan Ernst (Siemens)*

0930 Scaling Active Learning for the Development of Imagery-Derived Maps, *Ben Kadlec (Uber)*

1000 Morning Break

1030 Building Computer Vision Systems With Open Source Software, *Matt Turek (Kitware)*

1100 Transforming Research Code Into Robust Multiplatform Mobile Products, *Stephen Miller (Fyusion)*

1130 Transfer Learning: Data Curation, Training, and Deployment Strategies, *Tim Franklin (Microsoft)*

1200 Research to Prod: Large Scale Visual Recognition in the Cloud, *Wei Xia (Amazon)*

Inverse Reinforcement Learning for Computer Vision

Organizers: Nick Rhinehart
Paul Vernaza
Kris Kitani

Location: Ballroom B & D

Time: 0845-1200 (Half Day — Morning)

Description: Modern computer vision is great at analyzing pictures, but what about the people and robots in the pictures? What are their motivations? What will they do in the future? Are they behaving optimally? Inverse Reinforcement Learning (IRL) is well-suited to answer these kinds of questions. IRL provides a framework to learn and reason about the intent underlying goal-driven behavior. While there have been some successes in applying IRL to computer vision problems, we believe that IRL as a tool is underappreciated by computer vision researchers seeking to understand goal-driven behavior. In this tutorial, we will give an overview of the motivation, application, and practical aspects of IRL as applied to computer vision problems. The tutorial will be as self-contained as possible and will cover the relevant background material on Reinforcement Learning (RL).

Schedule:

0845 Introduction: What is IRL? What is RL?

0900 Successes of IRL

0915 A Reinforcement Learning Primer

1000 Morning Break

1030 Inverse Reinforcement Learning I

1100 Inverse Reinforcement Learning II

1130 Applications to Computer Vision In Real Life (IRL IRL)

Notes:

Differential Geometry for Engineers

Organizers: Jonathan Manton

Location: Ballroom H

Time: 0830-1230 (Half Day — Morning)

Description: This tutorial endeavours to introduce differential geometry simply, efficiently, rigorously and engagingly. A motivating example used throughout is optimisation on manifolds: how can gradient descent or a Newton method be performed on a manifold? An integral part of the tutorial is teaching how to do calculations involving manifolds, including how to compute derivatives, tangent spaces, curvature and so forth. At a higher level, the different ways of working with manifolds will be explained (e.g., extrinsic coordinates versus local coordinates), and ways for extending algorithms from Euclidean space to manifolds will be discussed.

Schedule:

0830 Motivation and Multi-Dimensional Differential Calculus

0915 Tangent Spaces and Directional Derivatives

1000 Morning Break

1030 Concrete Manifolds, Coordinate Charts and Smooth Functions

1115 Metrics, Parallel Transport and Connections

1200 Gradients, Hessians and Worked Examples

Notes:

Computational Imaging for Self-Driving Vehicles

Organizers: Jan Kautz
Ramesh Raskar
Achuta Kadambi
Guy Satat

Location: Room 355 A-C

Time: 1330-1730 (Half Day — Afternoon)

Description: To enable superhuman driving, we need superhuman vision. Here, we discuss recent advances in imaging hardware used in self-driving cars, and how this will impact the future of the industry. In this tutorial attendees will get an overview of the sensors that are equipped in self-driving vehicles. An emphasis will be placed on LIDAR, as the course practitioners believe this is the key instrument. Instructors will present recent papers in vision, graphics, and computational photography that exploit such sensors. Finally, attendees are provided a primer on how they can incorporate computational imaging into their own research.

Schedule:

1330 Computational Imaging and Implications to Self-Driving Vehicles, *Ramesh Raskar*

1415 Limitations With Existing Imaging Sensors on Cars, *Guy Satat, Achuta Kadambi*

1500 Q&A and Break

1510 Emerging Vision Sensors for Self-Driving Cars, *Achuta Kadambi*

1550 Imaging in Challenging Weather Conditions, *Guy Satat*

1630 Q&A and Break

1640 Deep Learning-Based Computational Imaging, *Jan Kautz*

1720 Discussion of Open Problems

Friday, June 22

0730–1500 Registration (North Lobby)

0730–0900 Breakfast (Halls 1-4)

1000–1100 Morning Break (Halls 1-4)

1200–1330 Lunch (Halls 1-4)

1530–1630 Afternoon Break (Halls 1-4)

Computer Vision for Physiological Measurement

Organizers: Wenjin Wang
Daniel McDuff
Sander Stuijk

Location: Room 250 B

Schedule: Full Day

0830 **Poster Setup** (all papers have poster panels for the whole day)

0850 **Opening Remarks**

0900 **Invited Talk:** *Steffen Leonhardt (RWTH Aachen Univ.)*

1000 Morning Break (Halls 1-4) **& Poster Session** (includes oral papers)

- Local Group Invariance for Heart Rate Estimation From Face Videos in the Wild, *Christian S. Pilz; Sebastian Zaunseder; Jarek Krajewski; Vladimir Blazek*
- Advertisement Effectiveness Estimation Based on Crowdsourced Multimodal Affective Responses, *Genki Okada; Kenta Masui; Norimichi Tsumura*
- SparsePPG: Towards Driver Monitoring Using Camera-Based Vital Signs Estimation in Near-Infrared, *Ewa Magdalena Nowara; Tim K. Marks; Hassan Mansour; Ashok Veeraraghavan*
- Novel Algorithms to Monitor Continuous Cardiac Activity With a Video Camera, *Gregory F. Lewis; Maria I. Davila; Stephen W. Porges*
- Measurement of Capillary Refill Time (CRT) in Healthy Subjects Using a Robotic Hand, *Emmett Kerr; Sonya Coleman; Martin McGinnity; Andrea Shepherd*
- A Novel Framework for Remote Photoplethysmography Pulse Extraction on Compressed Videos, *Changchen Zhao; Chun-Liang Lin; Weihai Chen; Zhengguo Li*
- Non-Contact Heart Rate Monitoring by Combining Convolutional Neural Network Skin Detection and Remote Photoplethysmography via a Low-Cost Camera, *Chuanxiang Tang; Jiwu Lu; Jie Liu*
- Exploring the Feasibility of Face Video Based Instantaneous Heart-Rate for Micro-Expression Spotting, *Puneet Gupta; Brojeshwar Bhowmick; Arpan Pal*
- Video Based Measurement of Heart Rate and Heart Rate Variability Spectrogram From Estimated Hemoglobin Information, *Munenori Fukunishi; Kouki Kurita; Shoji Yamamoto; Norimichi Tsumura*

1030 Periodic Variance Maximization Using Generalized Eigenvalue Decomposition Applied to Remote Photoplethysmography Estimation, *Richard Macwan; Serge Bobbia; Yannick Benezeth; Julien Dubois; Alamin Mansouri*

1055 Real-Time Temporal Superpixels for Unsupervised Remote Photoplethysmography, *Serge Bobbia; Duncan Luguern; Yannick Benezeth; Keisuke Nakamura; Randy Gomez; Julien Dubois*

1120 Fully-Automatic Camera-Based Pulse-Oximetry During Sleep, *Tom Vogels; Mark van Gastel; Wenjin Wang; Gerard de Haan*

1145 Lunch (Halls 1-4)

1400 **Invited Talk:** *William T. Freeman (MIT)*

1500 Impairing Factors in Remote-PPG Pulse Transit Time Measurements on the Face, *Andreia Moço; Sander Stuijk; Mark van Gastel; Gerard de Haan*

1525 Afternoon Break (Halls 1-4) **& Poster Session** (includes oral papers)

1555 Deep Super Resolution for Recovering Physiological Information From Videos, *Daniel McDuff*

1620 Direct-Global Separation for Improved Imaging Photoplethysmography, *Jaehee Park; Ashutosh Sabharwal; Ashok Veeraraghavan*

1645 **Best Paper Announcement & Closing Remarks**

Automated Analysis of Marine Video for Environmental Monitoring

Organizers: Anthony Hoogs
Benjamin L. Richards
David Kriegman

Location: Room 150 A-B

Schedule: Full Day

0900 **Welcome & Introduction**

0915 **Invited Talk:** Underwater Video Collection and Exploitation for Stock Assessment by the NOAA Fisheries Science Centers, *Randy Cutter (NOAA Southwest Fisheries Science Center)*

1000 Automated Analysis of Marine Video With Limited Data, *Deborah Levy; Yuval Belfer; Elad Osherov; Eyal Bigal; Aviad P. Scheinin; Hagai Nativ; Dan Tchernov; Tali Treibitz*

1030 Morning Break (Halls 1-4)

S1: Marine Flora and Fauna (1100-1245)

1100 **Invited Talk:** Dynamic Data Driven Observatories for Identification, *Sai Ravela (MIT)*

1145 A Comparison of Deep Learning Methods for Semantic Segmentation of Coral Reef Survey Images, *Andrew King; Suchendra M. Bhandarkar; Brian M. Hopkinson*

1215 Stingray Detection of Aerial Images Using Augmented Training Images Generated by a Conditional Generative Model, *Yi-Min Chou; Chien-Hung Chen; Keng-Hao Liu; Chu-Song Chen*

1245 Lunch (Halls 1-4)

S2: Marine Imaging and Measurement (1400-1600)

1400 **Invited Talk:** Statistical Tomography of Microscopic Life, *Yoav Schecter (Technion)*

1445 Camera Calibration for Underwater 3D Reconstruction Based on Ray Tracing Using Snell's Law, *Malte Pedersen; Stefan Hein Bengtson; Rikke Gade; Niels Madsen; Thomas B. Moeslund*

1515 **Invited Talk:** A Revised Underwater Image Formation Model, *Derya Akkaynak (Univ. of Haifa)*

1600 Afternoon Break (Halls 1-4)

S3: Marine Image Challenge Results (1615-1715)

1615 **Challenge Description**

1630 **Challenge Presentation —** First Place

1645 **Challenge Presentation —** Second Place

1700 **Challenge Presentation —** Third Place

Visual Odometry and Computer Vision Applications Based on Location Clues

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

Location: Room 257

Schedule: Full Day

0840 **Keynote Talk:** Localization in Urban Environments Using Single Images and Simple 2D Maps,
Vincent Lepetit (Univ. of Bordeaux & TU Graz)"

0920 **Keynote Talk:** TBA, *Manmohan Chandraker (NEC Labs & UCSD)*

1000 Drone-View Building Identification by Cross-View Visual Learning and Relative Spatial Estimation, *Chun-Wei Chen; Yin-Hsi Kuo; Tang Lee; Cheng-Han Lee; Winston Hsu*

1020 Integration of Absolute Orientation Measurements in the KinectFusion Reconstruction Pipeline, *Silvio Giancola; Jens Schneider; Peter Wonka; Bernard S. Ghanem*

1040 Morning Break (Halls 1-4)

1100 **Keynote Talk:** Vision Systems for Planetary Landers: Progress and Challenges, *Larry Matthies (NASA)*

1140 Optimal Linear Attitude Estimator for Alignment of Point Clouds, *Xue Iuan Wong; Taewook Lee; Puneet Singla; Manoranjan Majji*

1200 Multi-Scale Voxel Hashing and Efficient 3D Representation for Mobile Augmented Reality, *Yi Xu; Yuzhang Wu; Hui Zhou*

1220 Lunch (Halls 1-4) **& Poster Session**

1330 **Keynote Talk:** Unsupervised Learning of Depth and Ego-Motion using 3D Geometric Constraints, *Anelia Angelova (Google Brain)*

1410 **Keynote Talk:** TBA, *Ruigang Yang (Baidu Research & Univ. of Kentucky)*

1450 A Deep CNN-Based Framework for Enhanced Aerial Imagery Registration With Applications to UAV Geolocalization, *Ahmed Nassar; Karim Amer; Reda ElHakim; Mohamed ElHelw*

1510 Semantic Metric 3D Reconstruction for Concrete Inspection, *Liang Yang; Bing Li; Wei Li; Biao Jiang; Jizhong Xiao*

1530 Afternoon Break (Halls 1-4)

1550 **Keynote Talk:** Semantic Visual Localization, *Andreas Geiger (MPI & Univ. of Tübingen)*

1630 Deep Visual Teach and Repeat on Path Networks, *Tristan Swedish; Ramesh Raskar*

1650 Automated Virtual Navigation and Monocular Localization of Indoor Spaces From Videos, *Qiong Wu; Ambrose Li*

Efficient Deep Learning for Computer Vision

Organizers: Peter Vajda
Fernando De la Torre
Kurt Keutzer
Ramesh Sarukkai

Location: Ballroom F

Schedule: Full Day

0850 **Welcome**

0900 **Invited Talk:** TBA, *Ali Farhadi (Univ. of Washington)*

0930 **Invited Talk:** Co-Design of Deep Neural Nets and Neural Net Accelerators for Embedded Vision, *Kurt Keutzer (UC Berkeley)*

1000 Eye in the Sky: Real-Time Drone Surveillance System (DSS) for Violent Individuals Identification Using ScatterNet Hybrid Deep Learning Network, *Amarjot Singh; Devendra Patil; SN Omkar*

1010 SqueezeNext: Hardware-Aware Neural Network Design, *Amir Gholami; Kiseok Kwon; Bichen Wu; Zizheng Tai; Xiangyu Yue; Peter Jin; Sicheng Zhao; Kurt Keutzer*

1020 Recurrent Segmentation for Variable Computational Budgets, *Lane McIntosh; Niru Maheswaranathan; David Sussillo; Jonathon Shlens*

1030 Morning Break (Halls 1-4)

1100 **Invited Talk:** AMC: AutoML for Model Compression, *Song Han (Massachusetts Institute of Technology)*

1130 **Invited Talk:** TBA, *Forrest Iandola (DeepScale)*

1200 Highway Network Block With Gates Constraints for Training Very Deep Networks, *Oyebade K. Oyedotun; Abd El Rahman Shabayek; Djamila Aouada; Björn Ottersten*

1210 MUNet: Macro Unit-Based Convolutional Neural Network for Mobile Devices, *Dae Ha Kim; Seung Hyun Lee; Byung Cheol Song*

1220 Ultra Power-Efficient CNN Domain Specific Accelerator With 9.3TOPS/Watt for Mobile and Embedded Applications, *Baohua Sun; Lin Yang; Patrick Dong; Wenhan Zhang; Jason Dong; Charles Young*

1230 Lunch (Halls 1-4)

1300 **Invited Talk:** Quantization in Production, *Pete Warden (Google)*

1330 **Invited Talk:** TBA, *Jeff Gehlhaar (Qualcomm)*

1400 Merging Deep Neural Networks for Mobile Devices, *Yi-Min Chou; Yi-Ming Chan; Jia-Hong Lee; Chih-Yi Chiu; Chu-Song Chen*

1410 Efficient Deep Learning Inference Based on Model Compression, *Qing Zhang; Mengru Zhang; Mengdi Wang; Wanchen Sui; Chen Meng; Jun Yang; Weidan Kong; Xiaoyuan Cui; Wei Lin*

1420 Learning Network Architectures of Deep CNNs Under Resource Constraints, *Michael Chan; Daniel Scarafoni; Ronald Duarte; Jason Thornton; Luke Skelly*

1430 Afternoon Break (Halls 1-4) **& Poster Session**

1530 **Introduction & Presentations by the Panelists**

1630 **Panel:** High-Performance Training for Deep Learning and Computer Vision

- **Moderator:** *Ramesh Sarukkai*
- **Panelists:** *Michael James (Cerebras), Dave Driggers (Cirrascale), John Barrus (Google), Dhabaleswar Panda (Ohio State Univ.)*

1800 **Closing Remarks & Announcements**

ActivityNet Large Scale Activity Recognition

Organizers: Cees Snoek
Juan Carlos Niebles
Bernard Ghanem
Fabian Caba Heilbron
Humam Alwassel
Victor Escorcia
Ranjay Khrishna
Kenji Hata
Shyamal Buch
Cuong Duc Dao

Location: Room 260

Schedule: Full Day

0900 **Welcome & Opening Remarks**
0915 **Keynote Talk:** *Shih-Fu Chang (Columbia Univ.)*
1000 Morning Break (Halls 1-4)
1030 **Keynote Talk:** *Cordelia Schmid (INRIA)*
1115 **Panel Discussion:** Future of Activity Understanding
1200 **Description of All Tasks** (ActivityNet & Guest Tasks)
1230 Lunch (Halls 1-4)
1330 **Challenge Results:** ActivityNet Temporal Proposal Detection
1340 **Winner Presentation:** ActivityNet Temporal Proposal Detection
1400 **Challenge Results:** ActivityNet Temporal Activity Detection
1410 **Winner Presentation:** ActivityNet Temporal Activity Detection
1430 **Challenge Results:** ActivityNet Temporal Captioning
1440 **Winner Presentation:** ActivityNet Temporal Captioning
1500 Afternoon Break (Halls 1-4)
1530 **Challenge Results:** Kinetics
1540 **Winner Presentation:** Kinetics
1600 **Challenge Results:** AVA
1610 **Winner Presentation:** AVA
1630 **Challenge Results:** Moments in Time
1640 **Winner Presentation:** Moments in Time
1700 **Closing Remarks**

Joint Detection, Tracking, and Prediction in the Wild

Organizers: Amir Sadeghian
Agrim Gupta
Marynel Vazquez
Noriaki Hirose
Vineet Kosaraju
Michael Abbott
Alex Alahi
Silvio Savarese

Location: Room 250 E-F

Schedule: Full Day

1000 **Opening Remarks**
1015 **Invited Talk:** *Larry Davis (Univ. of Maryland)*
1100 **Invited Talk:** *Yaser Sheikh (Carnegie Mellon Univ.)*
1145 Spotlilght: DIY Human Action Dataset Generation, *Mehran Khodabandeh; Hamid Reza Vaezi Joze; Ilya Zharkov; Vivek Pradeep*
1153 Spotlilght: Convolutional Social Pooling for Vehicle Trajectory Prediction, *Nachiket Deo; Mohan M. Trivedi*
1200 Lunch (Halls 1-4)
1330 **Invited Talk:** *Rita Cucchiara (Univ. of Modena & Reggio Emilia)*
1415 **Invited Talk:** *Bernt Schiele (Max-Planck Institute)*
1500 Afternoon Break (Halls 1-4) **& Poster Session**

- HP-GAN: Probabilistic 3D Human Motion Prediction via GAN, *Emad Barsoum; John Kender; Zicheng Liu*
- Fusion of Head and Full-Body Detectors for Multi-Object Tracking, *Roberto Henschel; Laura Leal-Taixé; Daniel Cremers; Bodo Rosenhahn*
- Re-Identification for Online Person Tracking by Modeling Space-Time Continuum, *Neeti Narayan; Nishant Sankaran; Srirangaraj Setlur; Venu Govindaraju*
- DIY Human Action Dataset Generation, *Mehran Khodabandeh; Hamid Reza Vaezi Joze; Ilya Zharkov; Vivek Pradeep*
- Joint Detection and Online Multi-Object Tracking, *Hilke Kieritz; Wolfgang Hübner; Michael Arens*
- Convolutional Social Pooling for Vehicle Trajectory Prediction, *Nachiket Deo; Mohan M. Trivedi*

1600 **Invited Talk:** *Patrick Lucey (STATS)*

1645 **Panel**
1745 **Closing Remarks**
1830 **Dinner**

Computational Cameras and Displays

Organizers: Mohit Gupta
Sanjeev Koppal
Ioannis Gkioulekas

Location: Room 254 C

Schedule: Full Day

0830 **Welcome & Opening Remarks**
0845 **Keynote Talk:** Extra-Terrestrial Computational Imaging, With Down-to-Earth Outcomes, *Yoav Schechner (Technion - Israel Institute of Technology)*
0945 **Poster Spotlights**
1015 Morning Break (Halls 1-4)
1045 Space-Time-Brightness Sampling Using an Adaptive Pixel-Wise Coded Exposure, *Hajime Nagahara; Toshiki Sonoda; Dengyu Liu; Jinwei Gu*
1105 **Invited Talk:** TBA, *Guy Satat (MIT)*
1125 **Invited Talk:** TBA, *Emma Alexander (Harvard Univ.)*
1145 Lunch (Halls 1-4)
1315 Multi-Capture Dynamic Calibration of Multi-Camera Systems, *Avinash Kumar; Manjula Gururaj; Kalpana Seshadrinathan; Ramkumar Narayanswamy*
1335 Jittered Exposures for Image Super-Resolution, *Nianyi Li; Scott McCloskey; Jingyi Yu*
1355 **Invited Talk:** TBA, *Abe Davis (Stanford Univ.)*
1415 **Keynote Talk:** TBA, *Ori Katz (The Hebrew Univ. of Jerusalem)*
1515 Afternoon Break (Halls 1-4) **& Poster Session**
1615 **Keynote Talk:** TBA, *Laura Trutoiu (Magic Leap)*
1715 **Closing Remarks**

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: Room 259

Schedule: Full Day

0900 **Welcome**
0910 **Invited Talk:** *Jacob Steinhardt (Stanford Univ.)*

S1: Counterfeit and Forgery Detection (0950-1030)

0950 Discrete Cosine Transform Residual Feature Based Filtering Forgery and Splicing Detection in JPEG Images, *Aniket Roy; Diangarti Bhalang Tariang; Rajat Subhra Chakraborty; Ruchira Naskar*
1010 Forgery Detection in 3D-Sensor Images, *Noa Privman-Horesh; Azmi Haider; Hagit Hel-Or*
1030 Morning Break (Halls 1-4)
1100 **Invited Talk:** *Terrance Boult (Univ. of Colorado Colorado Springs)*
1140 **Poster Spotlights**

- Towards Reverse-Engineering Black-Box Neural Networks, *Seong Joon Oh; Max Augustin; Bernt Schiele; Mario Fritz*
- Generating Adversarial Images Using Genetic Algorithm, *Keeyoung Kim; Simon S Woo*
- Chaos Theory Meets Deep Learning: On Lyapunov Exponents and Adversarial Perturbations, *Vinay Uday Prabhu*
- (Almost) Data Agnostic Universal Adversarial Perturbations, *Valentin Khrulkov*
- On the Robustness of Semantic Segmentation Models to Adversarial Attacks, *Anurag Arnab; Ondrej Miksik; Philip Torr*
- On the Robustness of the CVPR 2018 White-Box Adversarial Example Defenses, *Anish Athalye; Nicholas Carlini*
- Boosting Adversarial Attacks With Momentum, *Yinpeng Dong; Fangzhou Liao; Tianyu Pang; Hang Su; Jun Zhu; Xiaolin Hu; Jianguo Li*

- Deflecting Adversarial Attacks With Pixel Deflection, *Aaditya Prakash; Nick Moran; Solomon Garber; Antonella DiLillo; James Storer*
- Robust Discriminative Localization Maps, *Aaditya Prakash; Nick Moran; Solomon Garber; Antonella DiLillo; James Storer*
- Robust Physical-World Attacks on Deep Learning Visual Classification, *Kevin Eykholt; Ivan Evtimov; Earlence Fernandes; Bo Li; Amir Rahmati; Chaiowei Xiao; Atul Prakash; Tadayoshi Kohno; Dawn Song*
- Art-Attack! On Style Transfers With Textures, Label Categories and Adversarial Examples, *Vinay Uday Prabhu*
- Seeing Voices and Hearing Faces: Cross-Modal Biometric Matching, *Samuel Albanie; Arsha Nagrani; Andrew Zisserman*
- Learning to Anonymize Faces for Privacy Preserving Action Detection, *Zhongzheng Ren; Yong Jae Lee; Michael Ryoo*
- Siamese Generative Adversarial Privatizer for Biometric Data, *Tomasz Trzcinski*
- Fully-Coupled Two-Stream Spatiotemporal Networks for Extremely Low Resolution Action Recognition, *Mingze Xu; Aidean Sharghi; Xin Chen; David Crandall*
- Fighting Fake News: Image Splice Detection via Learned Self-Consistency, *Jacob Huh; Andrew Liu; Andrew Owens; Alexei A. Efros*

1210 Lunch (Halls 1-4)

1330 **Invited Talk:** *Ian Goodfellow (Google Brain)*

S2: Privacy (1410-1510)

1410 VGAN-Based Image Representation Learning for Privacy-Preserving Facial Expression Recognition, *Jiawei Chen; Janusz Konrad; Prakash Ishwar*

1430 Privacy-Preserving Indoor Localization via Active Scene Illumination, *Jinyuan Zhao; Natalia Frumkin; Janusz Konrad; Prakash Ishwar*

1450 Human Perceptions of Sensitive Content in Photos, *Yifang Li; Wyatt Troutman; Bart P. Knijnenburg; Kelly Caine*

1510 Afternoon Break (Halls 1-4) **& Poster Session**

S3: Attacks Against Computer Vision Systems (1600-1720)

1600 On Visible Adversarial Perturbations & Digital Watermarking, *Jamie Hayes*

1620 On the Suitability of Lp-Norms for Creating and Preventing Adversarial Examples, *Mahmood Sharif; Lujo Bauer; Michael K. Reiter*

1640 Semantic Adversarial Examples, *Hossein Hosseini; Radha Poovendran*

1700 Convolutional Neural Networks for Iris Presentation Attack Detection: Toward Cross-Dataset and Cross-Sensor Generalization, *Steven Hoffman; Renu Sharma; Arun Ross*

1720 **Closing Remarks**

Bridges to 3D Vision

Organizers: David Fouhey
Qixing Huang
Joseph Lim
Hao Su
Shubham Tulsiani
David Forsyth
Jitendra Malik

Location: Room 251 A

Schedule: Full Day

Please visit bridgesto3d.github.io for final schedule!

0915 **Welcome / Overview**

0930 **Invited Talk:** *Leo Guibas (Stanford Univ.)*

1000 **Invited Talk:** *Kostas Daniilidis (Univ. of Pennsylvania)*

1030 Morning Break (Halls 1-4)

1100 **Invited Talk:** *Jitendra Malik (UC Berkeley)*

1130 **Invited Talk:** TBA

1200 Lunch (Halls 1-4)

1400 **Invited Talk:** TBA

1430 **Invited Talk:** TBA

1500 **Poster Session**

1530 Afternoon Break (Halls 1-4) **& Poster Session** (cont.)

1630 **Invited Talk:** *Thomas Funkhouser (Princeton Univ.)*

1700 **Panel Discussion**, *Invited Speakers*

Computer Vision in Sports

Organizers: Thomas B. Moeslund
Graham Thomas
Adrian Hilton
Peter Carr
Rikke Gade

Location: Room 250 C

Schedule: Full Day

0830 **Welcome**

0840 **Invited Talk:** TBA, *Irfan Essa (Georgia Tech)*

0925 SoccerNet: A Scalable Dataset for Action Spotting in Soccer Videos, *Silvio Giancola; Mohieddine Amine; Tarek Dghaily; Bernard Ghanem*

0945 Deep Decision Trees for Discriminative Dictionary Learning With Adversarial Multi-Agent Trajectories, *Tharindu Fernando; Sridha Sridharan; Clinton Fookes; Simon Denman*

1005 Morning Break (Halls 1-4)

1035 Part-Based Player Identification Using Deep Convolutional Representation and Multi-Scale Pooling, *Arda Senocak; Tae-Hyun Oh; Junsik Kim; In So Kweon*

1055 Fine-Grained Activity Recognition in Baseball Videos, *AJ Piergiovanni; Michael S. Ryoo*

1115 **Invited Talk:** Applications and Open Challenges for Computer Vision in High Performance Sport, *Stuart Morgan (La Trobe Univ.)*

1200 Lunch (Halls 1-4)

1330 **Invited Talk:** Watching With the Eye of an NBA Coach – Insights From Tracking a Full Season, *Ken Siebert, Andres Hasfura (Second Spectrum)*

1415 **Poster Spotlights**

1435 Afternoon Break (Halls 1-4) **& Poster Session**

- Soccer: Who Has the Ball? Generating Visual Analytics and Player Statistics, *Rajkumar Theagarajan; Federico Pala; Xiu Zhang; Bir Bhanu*
- Convolutional Neural Networks Based Ball Detection in Tennis Games, *Vito Renò; Nicola Mosca; Roberto Marani; Massimiliano Nitti; Tiziana D'Orazio; Ettore Stella*
- A Bottom-Up Approach Based on Semantics for the Interpretation of the Main Camera Stream in Soccer Games, *Anthony Cioppa; Adrien Deliège; Marc Van Droogenbroeck*
- Human Pose As Calibration Pattern; 3D Human Pose Estimation With Multiple Unsynchronized and Uncalibrated Cameras, *Kosuke Takahashi; Dan Mikami; Mariko Isogawa; Hideaki Kimata*
- Jersey Number Recognition With Semi-Supervised Spatial Transformer Network, *Gen Li; Shikun Xu; Xiang Liu; Lei Li; Changhu Wang*
- Kinematic Pose Rectification for Performance Analysis and Retrieval in Sports, *Dan Zecha; Moritz Einfalt; Christian Eggert; Rainer Lienhart*
- Automatic Cricket Highlight Generation Using Event-Driven and Excitement-Based Features, *Pushkar Shukla; Hemant Sadana; Apaar Bansal; Deepak Verma; Carlos Elmadjian; Balasubramanian Raman; Matthew Turk*
- Estimation of Center of Mass for Sports Scene Using Weighted Visual Hull, *Tomoya Kaichi; Shohei Mori; Hideo Saito; Kosuke Takahashi; Dan Mikami; Mariko Isogawa; Hideaki Kimata*
- A Directed Sparse Graphical Model for Multi-Target Tracking, *Mohib Ullah; Faouzi Alaya Cheikh*
- Estimating the Number of Soccer Players Using Simulation-Based Occlusion Handling, *Noor Ul Huda; Kasper H. Jensen; Rikke Gade; Thomas B. Moeslund*

1630 **Best Paper Award & Closing Remarks**

Fine-Grained Visual Categorization

Organizers: Ryan Farrell
Subhransu Maji
Yang Song
Xiao Zhang
Grant Van Horn
Oisin Mac Aodha
Yin Cui
David Rolnick

Location: Room 151 A-C

Schedule: Full Day

See schedule at: http://fgvc.org/FGVC5/

Women in Computer Vision

Organizers: Dena Bazazian
Ilke Demir
Adriana Romero
Viktoriia Sharmanska
Lyne P. Tchapmi

Location: Room 251 D

Schedule: Full Day

0850 **Welcome**

0900 **Keynote Talk:** *Jessica Hodgins (Carnegie Mellon Univ.)*

0930 **Keynote Talk:** *Laura Leal-Taixe (Technical Univ. Munich)*

1000 Morning Break (Halls 1-4) **& Poster Session**

- Autonomous Detection of Disruptions in the Intensive Care Unit Using Deep Mask R-CNN, *Kumar Rohit Malhotra; Anis Davoudi; Scott Siegel; Azra Bihorac; Parisa Rashidi*
- Encapsulating the Impact of Transfer Learning, Domain Knowledge and Training Strategies in Deep-Learning Based Architecture: A Biometric Based Case Study, *Avantika Singh; Aditya Nigam*
- Cross-Domain Fashion Image Retrieval, *Bojana Gajić; Ramón Baldrich*
- Word Spotting in Scene Images Based on Character Recognition, *Dena Bazazian; Dimosthenis Karatzas; Andrew D. Bagdanov*
- A Holistic Framework for Addressing the World Using Machine Learning, *Ilke Demir; Forest Hughes; Aman Raj; Kaunil Dhruv; Suryanarayana Murthy Muddala; Sanyam Garg; Barrett Doo; Ramesh Raskar*
- I Know How You Feel: Emotion Recognition With Facial Landmarks, *Ivona Tautkute; Tomasz Trzcinski; Adam Bielski*
- Early Diagnosis of Alzheimer's Disease: A Neuroimaging Study With Deep Learning Architectures, *Jyoti Islam; Yanqing Zhang*
- Cosmetic Features Extraction by a Single Image Makeup Decomposition, *Kanami Yamagishi; Shintaro Yamamoto; Takuya Kato; Shigeo Morishima*
- Automatic Large-Scale 3D Building Shape Refinement Using Conditional Generative Adversarial Networks, *Ksenia Bittner; Marco Körner*
- SAM: Pushing the Limits of Saliency Prediction Models, *Marcella Cornia; Lorenzo Baraldi; Giuseppe Serra; Rita Cucchiara*
- RPIfield: A New Dataset for Temporally Evaluating Person Re-Identification, *Meng Zheng; Srikrishna Karanam; Richard J. Radke*
- Large-Scale Ecological Analyses of Animals in the Wild Using Computer Vision, *Mikayla Timm; Subhransu Maji; Todd Fuller*
- Discovering Style Trends Through Deep Visually Aware Latent Item Embeddings, *Murium Iqbal; Adair Kovac; Kamelia Aryafar*
- Towards More Accurate Radio Telescope Images, *Nezihe Merve Gürel; Paul Hurley; Matthieu Simeoni*
- Multi-Resolution Deep Object Recognition Network, *Akram Bayat; Marc Pomplun*
- Cross-Modal Embeddings for Video and Audio Retrieval, *Amanda Duarte; Didac Surís; Amaia Salvador; Jordi Torres; Xavier Giró-i-Nieto*
- A Revised Underwater Image Formation Model, *Derya Akkaynak; Tali Treibitz*
- Variational Depth and Normal Fusion Algorithms for 3D Reconstruction, *Doris Antensteiner; Svorad Štolc; Thomas Pock*
- fMVR:Shall I Get My Video Back? Feature Matching Based Video Reconstruction, *Ekta Gujral*
- Image Segmentation Using Sparse Subset Selection, *Fariba Zohrizadeh; Farhad Kamangar*
- GANerated Hands for Real-Time 3D Hand Tracking From Monocular RGB, *Franziska Mueller; Florian Bernard; Oleksandr Sotnychenko; Dushyant Mehta; Srinath Sridhar; Dan Casas; Christian Theobalt*
- Do Hashtags Help? – Image Aesthetics Prediction Using Only Hashtags, *Hiya Roy; Toshihiko Yamasaki; Tatsuaki Hashimoto*
- Dynamic Facial Expression Recognition Through Visual Rhythm and Motion History Image, *Jadisha Yarif Ramírez Cornejo; Helio Pedrini*
- Deep Saliency Detection: From Supervised Learning to Unsupervised Learning, *Jing Zhang; Shaodi You; Yuchao Dai*
- Recognizing Elevator Buttons and Labels for Blind Navigation, *Jingya Liu; Yingli Tian*

- Keep It Short: Understanding Traffic Scenes From Very Short Representations, *Karla Brkić; Ivan Sikirić*
- Netizen-Style Commenting on Fashion Photos: Dataset and Diversity Measures, *Kuan-Ting Chen; Wen Hua Lin; Hung Yueh Chiang; Winston Hsu*
- Multi-Scale Spatially Weighted Local Histogram in O(1), *Mahdieh Poostchi; Ali Shafiekhani; George Thoma*
- Malaria Parasite Detection and Quantification Using Deep Neural Network, *Mahdieh Poostchi; Zhaohui Liang; Stefan Jaeger; George Thoma*
- A Neuro-Tensorial Approach for Learning Disentangled Representations, *Mengjiao Wang; Zhixin Shu; Shiyang Cheng; Yannis Panagakis; Dimitris Samaras; Stefanos Zafeiriou*
- I-ME: Iterative Model Evolution for Learning Activities From Weakly Labeled Videos, *Ozge Yalcinkaya; Eren Golge; Pinar Duygulu*
- View-Graph Selection Framework for SfM, *Rajvi Shah; P. J. Narayanan*
- Wearable RGB Camera-Based Navigation System for the Visually Impaired, *Reham Abobeah; Mohamed Hussein*
- UG^2: A Video Benchmark for Assessing the Impact of Image Restoration and Enhancement on Automatic Visual Recognition, *Rosaura G. Vidal; Sreya Banerjee; Klemen Grm; Vitomir Štruc; Walter J. Scheirer*
- Net2Vec: Explaining How Concepts Are Encoded in Deep Neural Networks, *Ruth Fong; Andrea Vedaldi*
- Activity Recognition Under Energy Constraints, *Sheila Pinto Caceres; Rafael Possas; Fabio Ramos*
- PhaseNet for Video Frame Interpolation, *Simone Meyer; Abdelaziz Djelouah; Brian McWilliams; Alexander Sorkine-Hornung; Markus Gross; Christopher Schroers*
- Deformation Aware Image Compression, *Tamar Rott Shaham; Tomer Michaeli*
- Hybridization of Feature Selection Methods Based on Tournament Design: HEp-2 Cell Image Classification, *Vibha Gupta; Arnav Bhavsar*
- Quick Adaptation of Segmentation FCN via Network Modulation, *Yanran Wang*
- To Be Focused: Efficient Weakly Supervised Learning via Soft Proposal Network, *Yi Zhu; Yanzhao Zhou; Qixiang Ye; Jianbin Jiao*
- Learning Peak Response for Weakly Supervised Instance-Level Segmentation, *Yi Zhu; Yanzhao Zhou; Qixiang Ye; Jianbin Jiao*

1130 On the Iterative Refinement of Densely Connected Representation Levels for Semantic Segmentation, *Arantxa Casanova; Guillem Cucurull; Michal Drozdzal; Adriana Romero; Yoshua Bengio*

1150 ARC: Adversarial Robust Cuts for Semi-Supervised and Multi-Label Classification, *Sima Behpour; Wei Xing; Brian D. Ziebart*

1210 Lunch (Halls 1-4)

1330 **Invited Talk:** *Octavia Camps (Northeastern Univ.)*

1400 Don't Just Assume; Look and Answer: Overcoming Priors for Visual Question Answering, *Aishwarya Agrawal; Dhruv Batra; Devi Parikh; Aniruddha Kembhavi*

1420 Learnable PINs: Cross-Modal Embeddings for Person Identity, *Arsha Nagrani; Samuel Albanie; Andrew Zisserman*

1450 **Invited Talk:** *Carol E. Reiley (drive.ai)*

1520 Gradient-Free Policy Architecture Search and Adaptation, *Sayna Ebrahimi; Anna Rohrbach; Trevor Darrell*

1540 Joint Event Detection and Description in Continuous Video Streams, *Huijuan Xu; Boyang Li; Vasili Ramanishka; Leonid Sigal; Kate Saenko*

1600 Afternoon Break (Halls 1-4)

1630 **Panel**, *Jessica Hodgins (Carnegie Mellon Univ.); Laura Leal-Taixe (Technical Univ. Munich); Timnit Gebru (Microsoft); Carol Reiley (drive.ai); Octavia Camps (Northeastern Univ.)*

1710 **Closing Remarks**

1900 **WiCV Mentorship Banquet** (On Thursday evening, June 21)

Notes:

Analysis and Modeling of Faces and Gestures

Organizers: Thomas S. Huang
Yun Raymond Fu
Matthew A. Turk
Michael Jones
Ming Shao
Joseph P. Robinson

Location: Room 250 D

Schedule: Full Day

0830 **Invited Talk:** *Jan Kautz (NVIDIA)*

0910 Recognizing American Sign Language Gestures From Within Continuous Videos, *Yuancheng Ye; Yingli Tian; Matt Huenerfauth; Jingya Liu*

0930 Fine-Grained Head Pose Estimation Without Keypoints, *Nataniel Ruiz; Eunji Chong; James M. Rehg*

0950 Generative Adversarial Style Transfer Networks for Face Aging, *Sveinn Pálsson; Eirikur Agustsson; Radu Timofte; Luc Van Gool*

1010 Empirically Analyzing the Effect of Dataset Biases on Deep Face Recognition Systems, *Adam Kortylewski; Bernhard Egger; Andreas Schneider; Thomas Gerig; Andreas Morel-Forster; Thomas Vetter*

1030 Morning Break (Halls 1-4)

1100 Motion Fused Frames: Data Level Fusion Strategy for Hand Gesture Recognition, *Okan Köpüklü; Neslihan Köse; Gerhard Rigoll*

1120 Clothing Change Aware Person Identification, *Jia Xue; Zibo Meng; Karthik Katipally; Haibo Wang; Kees van Zon*

1140 A Compact Deep Learning Model for Robust Facial Expression Recognition, *Chieh-Ming Kuo; Shang-Hong Lai; Michel Sarkis*

1200 Lunch (Halls 1-4)

1330 **Invited Talk:** *Jianchao Yang (Snapchat)*

1410 FACSCaps: Pose-Independent Facial Action Coding With Capsules, *Itir Onal Ertugrul; László A. Jeni; Jeffrey F. Cohn*

1430 Unraveling Human Perception of Facial Aging Using Eye Gaze, *Daksha Yadav; Naman Kohli; Ekampreet Kalsi; Mayank Vatsa; Richa Singh; Afzel Noore*

1450 Improving Viseme Recognition Using GAN-Based Frontal View Mapping, *Dário Augusto Borges Oliveira; Andrea Britto Mattos; Edmilson da Silva Morais*

1510 Light-Weight Head Pose Invariant Gaze Tracking, *Rajeev Ranjan; Shalini De Mello; Jan Kautz*

1530 Afternoon Break (Halls 1-4)

1600 **Invited Talk:** *Ajmal Mian (Univ. of Western Australia)*

1640 Implementing a Robust Explanatory Bias in a Person Re-Identification Network, *Esube Bekele; Wallace E. Lawson; Zachary Horne; Sangeet Khemlani*

1700 On Detecting Domestic Abuse via Faces, *Puspita Majumdar; Saheb Chhabra; Richa Singh; Mayank Vatsa*

1720 **Closing Remarks**

Towards Automatic Understanding of Visual Advertisements

Organizers: Adriana Kovashka
James Hahn

Location: Room 150 D-F

Schedule: Full Day

0900 **Welcome & Brainstorming**

0930 **Invited Talk:** TBA, *Jiebo Luo (Univ. of Rochester)*

1000 **Invited Talk:** Applications of Computer Vision in Display and Video Advertising, *Jesse Berent (Google)*

1030 Morning Break (Halls 1-4)

1100 **Brainstorming**

1130 **Challenge Description & Winner Talk**

1200 Lunch (Halls 1-4)

1330 **Invited Talk:** Decoding Political Advertising by Deep Learning, *Jungseock Joo (UCLA)*

1400 **Invited Talk:** Creating Visual Metaphors With Crowds and Machines, *Lydia Chilton (Columbia Univ.)*

1430 **Invited Talk:** Modelling Metaphor With Linguistic and Visual Features, *Ekaterina Shutova (Univ. of Amsterdam)*

1500 Afternoon Break (Halls 1-4) **& Poster Session**

1630 **Brainstorming & Closing Remarks**

Computer Vision for Microscopy Image Analysis

Organizers: Mei Chen
Dimitris Metaxas
Peter Bajcsy
Daniel Hoeppner

Location: Room 252

Schedule: Full Day

0900 **Welcome**

0915 **Keynote Talk:** Applications of Deep Learning to Critical Obstacles Understanding and Treating Neurodegenerative Disease, *Steve Finkbeiner (Univ. of California, San Francisco)*

1015 Morning Break (Halls 1-4)

S1: Spotlights of Accepted Papers 1 (1030-1130)

1030 Cell Image Segmentation by Integrating Multiple CNNs, *Yuki Hiramatsu; Kazuhiro Hotta; Ayako Imanishi; Michiyuki Matsuda; Kenta Terai*

1040 Large Kernel Refine Fusion Net for Neuron Membrane Segmentation, *Dongnan Liu; Donghao Zhang; Yang Song; Chaoyi Zhang; Heng Huang; Mei Chen; Weidong Cai*

1050 Three Dimensional Fluorescence Microscopy Image Synthesis and Segmentation, *Chichen Fu; Soonam Lee; David Joon Ho; Shuo Han; Paul Salama; Kenneth W. Dunn; Edward J. Delp*

1100 Improved Extraction of Objects From Urine Microscopy Images With Unsupervised Thresholding and Supervised U-Net Techniques, *Abdul Aziz; Harshit Pande; Bharath Cheluvaraju; Tathagato Rai Dastidar*

1110 Multilayer Encoder-Decoder Network for 3D Nuclear Segmentation in Spheroid Models of Human Mammary Epithelial Cell Lines, *Mina Khoshdeli; Garrett Winkelmaier; Bahram Parvin*

1120 Resolution-Enhanced Lensless Color Shadow Imaging Microscopy Based on Large Field-of-View Submicron-Pixel Imaging Sensors, *Cheng Yang; Haowen Ma; Xu Cao; Xia Hua; Xiaofeng Bu; Limin Zhang; Tao Yue; Feng Yan*

1130 **Panel:** Industry/Government Research Priority & Funding Opportunities, *Jie Yang (NSF), Michelle Freund (NIH), Peter Bajcsy (NIST), TBD (Merck)*

1230 Lunch (Halls 1-4)

1330 **Invited Talk:** Assessment of Intra-Tumor Heterogeneity and Evolutionary Trajectories in Cancer, *Subhajyoti De (Rutgers Univ.)*

1400 **Invited Talk:** Mapping the Brain at Cellular Resolution: Academic and Pharmaceutical Industry Applications, *Pavel Olsen (Cold Spring Habor Laboratory)*

S2: Spotlights of Accepted Papers 2 (1430-1530)

1430 Sequential Modeling of Deep Features for Breast Cancer Histopathological Image Classification, *Vibha Gupta; Arnav Bhavsar*

1440 Comparison of Deep Transfer Learning Strategies for Digital Pathology, *Romain Mormont; Pierre Geurts; Raphaël Marée*

1450 3D Cell Nuclear Morphology: Microscopy Imaging Dataset and Voxel-Based Morphometry Classification Results, *Alexandr A. Kalinin; Ari Allyn-Feuer; Alex Ade; Gordon-Victor Fon; Walter Meixner; David Dilworth; Jeffrey R. de Wet; Gerald A. Higgins; Gen Zheng; Amy Creekmore; John W. Wiley; James E. Verdone; Robert W. Veltri; Kenneth J. Pienta; Donald S. Coffey; Brian D. Athey; Ivo D. Dinov*

1500 FastSME: Faster and Smoother Manifold Extraction From 3D Stack, *Sreetama Basu; Elton Rexhepaj; Nathalie Spassky; Auguste Genovesio; Rasmus Reinhold Paulsen; ASM Shihavuddin*

1510 Localization and Tracking in 4D Fluorescence Microscopy Imagery, *Shahira Abousamra; Shai Adar; Natalie Elia; Roy Shilkrot*

1520 Estimation of Sperm Concentration and Total Motility From Microscopic Videos of Human Semen Samples, *Karan Dewan; Tathagato Rai Dastidar; Maroof Ahmad*

1530 Afternoon Break (Halls 1-4) **& Poster Session**

- Impact of Sampling and Augmentation on Generalization Accuracy of Microscopy Image Segmentation Methods, *Joe Chalfoun; Mary Brady; Michael Majurski; Peter Bajcsy; Petru Manesc*
- Understanding Pixel-to-Label Prediction Model, *Petru Manescu; Nathan Hotaling; Carl Simon; Sarala Padi; Peter Bajcsy; Nicholas Schaub*
- Cellular Structure Segmentation in 3D Microscopy Images, *Susanne Rafelski; Liya Ding; Jianxu Chen*

1630 **Invited Talk:** Life and Death Decisions—Classification, Characterization and Predictions of Death in Neuronal Models of Neurodegenerative Disease, *Jeremy Linsle (Univ. of California, San Francisco)*

1700 **Invited Talk:** Efficient and Scalable Tools for Large Scale Microscopy Image Analysis, *Erhan Bas (Janelia Research Campus, HHMI)*

1730 **Closing Remarks**

Mutual Benefits of Cognitive and Computer Vision: How Can We Use One to Understand the Other?

Organizers: Ali Borji
Krista A. Ehinger
Odelia Schwartz
Gregory Zelinsky
Hamed R.Tavakoli

Location: Room 250 A

Schedule: Full Day

0845 **Welcome**

0900 **Invited Talk:** What Constitutes a Theory of Vision? *James Elder (York Univ.)*

0950 **Invited Talk:** Predicting Goal-Directed Attention Control: A Tale of Three Deep Networks, *Gregory Zelinsky (Stony Brook)*

1035 Morning Break (Halls 1-4) **& Posters**

- ViS-HuD: Using Visual Saliency to Improve Human Detection With Convolutional Neural Networks, *Vandit Gajjar; Yash Khandhediya; Ayesha Gurnani; Viraj Mavani; Mehul S. Raval*
- Learning Biomimetic Perception for Human Sensorimotor Control, *Masaki Nakada; Honglin Chen; Demetri Terzopoulos*
- Assessing Shape Bias Property of Convolutional Neural Networks, *Hossein Hosseini; Baicen Xiao; Mayoore Jaiswal; Radha Poovendran*
- Deep-BCN: Deep Networks Meet Biased Competition to Create a Brain-Inspired Model of Attention Control, *Hossein Adeli; Gregory Zelinsky*
- Image Caption Generation With Hierarchical Contextual Visual Spatial Attention, *Mahmoud Khademi; Oliver Schulte*
- Estimating Attention of Faces Due to Its Growing Level of Emotions, *Ravi Kant Kumar; Jogendra Garain; Dakshina Ranjan Kisku; Goutam Sanyal*
- Totally Looks Like – How Humans Compare, Compared to Machines, *Amir Rosenfeld; Markus D. Solbach; John K. Tsotsos*
- Fusing Visual Saliency for Material Recognition, *Lin Qi; Ying Xu; Xiaowei Shang; Junyu Dong*
- Increasing Video Saliency Model Generalizability by Training for Smooth Pursuit Prediction, *Mikhail Startsev; Michael Dorr*
- Representation of Categories in Filters of Deep Neural Networks, *Katerina Malakhova*

1200 Lunch (Halls 1-4)

1400 **Invited Talk:** Studying Visual Learning in Children Using Computer Vision, *David Crandall (Indiana Univ.)*

1450 Using Psychophysical Methods to Understand Mechanisms of Face Identification in a Deep Neural Network, *Tian Xu; Oliver Garrod; Steven H. Scholte; Robin Ince; Philippe G. Schyns*

1510 Relating Deep Neural Network Representations to EEG-fMRI Spatiotemporal Dynamics in a Perceptual Decision-Making Task, *Tao Tu; Jonathan Koss; Paul Sajda*

1530 Afternoon Break (Halls 1-4)

1600 Scene Grammar in Human and Machine Recognition of Objects and Scenes, *Akram Bayat; Do Hyong Koh; Anubhaw Kumar Nand; Marta Pereira; Marc Pomplun*

1620 Audio-Visual Temporal Saliency Modeling Validated by fMRI Data, *Petros Koutras; Georgia Panagiotaropoulou; Antigoni Tsiami; Petros Maragos*

1640 Priming Neural Networks, *Amir Rosenfeld; Mahdi Biparva; John K. Tsotsos*

1700 **Panel Discussion**

Vision With Biased or Scarce Data

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

Location: Room 251 B

Schedule: Full Day

0845 **Welcome & Introductory Remarks**

0910 **Keynote Talk:** Human-Machine Collaboration for Large-Scale Image Annotation, *Vittorio Ferrari (Google Research)*

0945 **Keynote Talk:** Making Your Data Count: Sharing Information Across Domains and Tasks, *Judy Hoffman (UC Berkeley)*

1020 Morning Break (Halls 1-4) **& Poster Session**

1050 Markov Chain Neural Networks, *Maren Awiszus; Bodo Rosenhahn*

1110 A Generative Model for Zero Shot Learning Using Conditional Variational Autoencoders, *Ashish Mishra; Shiva Krishna Reddy; Anurag Mittal; Hema A. Murthy*

1130 Endoscope Navigation and 3D Reconstruction of Oral Cavity by Visual SLAM With Mitigated Data Scarcity, *Liang Qiu; Hongliang Ren*

1150 Lunch (Halls 1-4)

1320 **Keynote Talk:** Sparsity and Scarcity in the Automotive Domain, *Jonas Uhrig (Daimler AG & Univ. of Freiburg)*

1355 **Keynote Talk:** How to Satisfy the Thirst for Data? *Andreas Geiger (Univ. of Tübingen & MPI-IS)*

1430 Detecting Anomalous Faces With "No Peeking" Autoencoders, *Anand Bhattad, Jason Rock, David Forsyth*

1450 Afternoon Break (Halls 1-4) **& Poster Session**

1530 **Keynote Talk:** Three Approaches to Training Object and Activity Detectors With Less Annotation, *Jeff Siskind (Purdue Univ.)*

1605 **Keynote Talk:** Fairness in Computer Vision, *Olga Russakovsky (Princeton Univ.)*

1640 **Concluding Remarks**

Beyond Supervised Learning

Organizers: Amir R. Zamir
Jitendra Malik
Alexei Efros
Leonidas Guibas
Josh Tenenbaum
Silvio Savarese

Location: Ballroom G

Schedule: Full Day

Confirmed Invited Talks: See final schedule at http://beyond-supervised.ai/ for times & full list of speakers.

- **Invited Talk:** Forcing Vision and Language Models to Not Just Talk but Also Actually See, *Devi Parikh (Georgia Tech; FAIR)*
- **Invited Talk:** TBA, *Honglak Lee (U. of Michigan; Google)*
- **Invited Talk:** Learning From Unlabeled Video, *Virginia De Sa (UCSD)*
- **Invited Talk:** TBA, *Carl Vondrick (Columbia, Google)*
- **Invited Talk:** TBA, *Dan Yamins (Stanford Univ.)*

0900 **Welcome**

0910 **Talks:** Session 1

1000 Morning Break (Halls 1-4)

1030 **Talks:** Session 2

1200 Lunch (Halls 1-4)

1400 **Talks:** Session 3

1530 Afternoon Break (Halls 1-4)

1600 **Talks:** Session 4

1730 **Dinner**

Notes:

Real World Challenges and New Benchmarks for Deep Learning in Robotic Vision

Organizers: Niko Sünderhauf
Anelia Angelova
Gustavo Carneiro
Kevin Murphy
Anton van den Hengel
Vijay Kumar
Jürgen Leitner
Trung T. Pham
Feras Dayoub
Ingmar Posner
Michael Milford
Ian Reid
Peter Corke

Location: Room 150 G

Schedule: Full Day

0900 **Welcome**

0920 **Invited Talk:** *Andreas Geiger (Univ. of Tübingen, MPI)*

0940 **Poster Spotlights**

1000 Morning Break (Halls 1-4) **& Posters**

1030 **Invited Talk:** *Oliver Brock (TU Berlin)*

1050 **Invited Talk:** *Walter Scheirer (Univ. of Notre Dame)*

1110 **Invited Talk:** *Dieter Fox (Univ. of Washington)*

1130 **Discussion: Research Challenges in Robotic Vision**

1200 Lunch (Halls 1-4)

1400 **New Robotic Vision Benchmarks**

1420 **Invited Talk:** *Larry Zitnick (Facebook AI Research)*

1440 **Invited Talk:** *Kristen Grauman (Univ. of Texas Austin)*

1500 **Poster Session**

- VisDA: A Synthetic-to-Real Benchmark for Visual Domain Adaptation, *Xingchao Peng; Ben Usman; Neela Kaushik; Dequan Wang; Judy Hoffman; Kate Saenko*
- Paris-Lille-3D: A Point Cloud Dataset for Urban Scene Segmentation and Classification, *Xavier Roynard; Jean-Emmanuel Deschaud; François Goulette*
- New Metrics and Experimental Paradigms for Continual Learning, *Tyler L. Hayes; Ronald Kemker; Nathan D. Cahill; Christopher Kanan*
- Action-Conditioned Convolutional Future Regression Models for Robot Imitation Learning, *Alan Wu; AJ Piergiovanni; Michael S. Ryoo*
- Falling Things: A Synthetic Dataset for 3D Object Detection and Pose Estimation, *Jonathan Tremblay; Thang To; Stan Birchfield*
- Learning Instance Segmentation by Interaction, *Deepak Pathak; Yide Shentu; Dian Chen; Pulkit Agrawal; Trevor Darrell; Sergey Levine; Jitendra Malik*
- Active Vision Dataset Benchmark, *Phil Ammirato; Alexander C. Berg; Jana Košecká*
- Zero-Shot Visual Imitation, *Deepak Pathak; Parsa Mahmoudieh; Guanghao Luo; Pulkit Agrawal; Dian Chen; Yide Shentu; Evan Shelhamer; Jitendra Malik; Alexei A. Efros; Trevor Darrell*
- Embodied Question Answering, *Abhishek Das; Samyak Datta; Georgia Gkioxari; Stefan Lee; Devi Parikh; Dhruv Batra*

1530 Afternoon Break (Halls 1-4)

1600 **Invited Talk:** *Vladen Koltun (Intel Intelligent Systems Lab)*

1620 **Invited Talk:** *Dhruv Batra (Georgia Tech)*

1640 **Panel Discussion & Closing Remarks**

Visual Understanding of Subjective Attributes of Data

Organizers: Xavier Alameda-Pineda
Miriam Redi
Adriana Kovashka
Devi Parikh
Nicu Sebe
Shih-Fu Chang

Location: Room 254 A

Schedule: Half Day - Morning

0830 **Welcome**

0840 **Invited Talk:** *Antonio Torralba (MIT)*

0920 Human Action Adverb Recognition: ADHA Dataset and a Three-Stream Hybrid Model, *Bo Pang; Kaiwen Zha; Cewu Lu*

0935 Pay Attention to Virality: Understanding Popularity of Social Media Videos With the Attention Mechanism, *Adam Bielski; Tomasz Trzcinski*

0950 **Invited Talk:** *Kristen Grauman (UT Austin)*

1030 Morning Break (Halls 1-4) **& Poster Session**

- Learning Fashion by Simulated Human Supervision, *Eli Alshan; Sharon Alpert; Assaf Neuberger; Nathaniel Bubis; Eduard Oks*
- Finding Your Lookalike: Measuring Face Similarity Rather Than Face Identity, *Amir Sadovnik; Wassim Gharbi; Thanh Vu; Andrew Gallagher*
- Behavior and Personality Analysis in a Nonsocial Context Dataset, *Dario Dotti; Mirela Popa; Stylianos Asteriadis*
- Ambiance in Social Media Venues: Visual Cue Interpretation by Machines and Crowds, *Gülcan Can; Yassir Benkhedda; Daniel Gatica-Perez*
- From Apparent to Real Age: Gender, Age, Ethnic, Makeup, and Expression Bias Analysis in Real Age Estimation, *Albert Clapés; Ozan Bilici; Dariia Temirova; Egils Avots; Gholamreza Anbarjafari; Sergio Escalera*

1130 **Invited Talk:** Apparent Human Behavior Understanding, *Sergio Escalera (CVC)*

1210 **Closing Remarks**

Computational Models for Learning Systems and Educational Assessment

Organizers: Saad M. Khan
Yuchi Huang
Terrance E. Boult
Alina von Davier

Location: Room 150 C

Schedule: Half Day - Morning

0845 **Welcome**

0900 **Invited Talk:** *Marcelo Worsley (Northwestern Univ.)*

0945 **Invited Talk:** *Bror Saxberg (Chan Zuckerberg Initiative)*

1030 Morning Break (Halls 1-4)

1040 Scaling Handwritten Student Assessments With a Document Image Workflow System, *Vijay Rowtula; Varun Bhargavan; Mohan Kumar; C.V. Jawahar*

1100 Teachers' Perception in the Classroom, *Ömer Sümer; Patricia Goldberg; Kathleen Stürmer; Tina Seidel; Peter Gerjets; Ulrich Trautwein; Enkelejda Kasneci*

1120 **Invited Talk:** *Ramona Pierson (Amazon)*

1200 **Closing Remarks**

Autonomous Driving

Organizers: Jose M. Alvarez
Fisher Yu
Ruigang Yang
Antonio M. Lopez
Andreas Geiger
Alan Yuille
Dinesh Manocha
Dequan Wang
David Vázquez
Hongdong Li
John Leonard
Jianxiong Xiao
Markus Enzweiler
Trevor Darrell
Tomas Pajdla

Location: Room 251 E-F

Schedule: Half Day - Morning

0850 **Opening Remarks**

0900 **Invited Talk:** TBA, *David A. Forsyth (Univ. of Illinois at Urbana-Champaign)*

0940 **Challenge Awards**

0950 **Challenge Winner Presentation —** Task 1: Drivable Area Segmentation

1000 **Challenge Winner Presentation —** Task 2: Road Object Detection

1010 **Challenge Winner Presentation —** Task 3: Domain Adaption of Semantic Segmentation

1020 **Challenge Winner Presentation —** Task 4: Instance-level Video Segmentation

1030 Morning Break (Halls 1-4) **& Poster Session** (top 3 papers in each task)

1130 **Invited Talk:** Mapping the world for Autonomous – Opportunities and Challenges, *Ramesh Sarukkai (Lyft)*

1210 **Closing Remarks**

Sight and Sound

Organizers: Andrew Owens
Jiajun Wu
William T. Freeman
Andrew Zisserman
Jean-Charles Bazin
Zhengyou Zhang
Antonio Torralba
Alexei Efros

Location: Room 251 C

Schedule: Half Day - Afternoon

1330 **Welcome**

1335 Paper Session 1

1400 **Invited Talk:** *Antonio Torralba (MIT)*

1430 **Invited Talk:** *Joon Son Chung (Oxford)*

1500 Paper Session 2

1530 Afternoon Break (Halls 1-4) **& Posters**

- Learning to Separate Object Sounds by Watching Unlabeled Video, *Ruohan Gao; Rogerio S. Feris; Kristen Grauman*
- Visual to Sound: Generating Natural Sound for Videos in the Wild, *Yipin Zhou; Zhaowen Wang; Chen Fang; Trung Bui; Tamara L. Berg*
- Fast Forwarding Egocentric Videos by Listening and Watching, *Vinicius S, Furlan; Ruzena Bajcsy; Erickson R. Nascimento*
- Learnable PINs: Cross-Modal Embeddings for Person Identity, *Arsha Nagrani; Samuel Albanie; Andrew Zisserman*
- The Sound of Pixels, *Hang Zhao; Chuang Gan; Andrew Rouditchenko; Carl Vondrick; Josh McDermott; Antonio Torralba*
- On Learning Association of Sound Source and Visual Scenes, *Arda Senocak; Tae-Hyun Oh; Junsik Kim; Ming-Hsuan Yang; In So Kweon*
- Image Generation Associated With Music Data, *Yue Qiu; Hirokatsu Kataoka*
- Semantic Speech Retrieval With a Visually Grounded Model of Untranscribed Speech, *Herman Kamper; Gregory Shakhnarovich; Karen Livescu*
- Weakly Supervised Representation Learning for Unsynchronized Audio-Visual Events, *Sanjeel Parekh; Slim Essid; Alexey Ozerov; Ngoc Q. K. Duong; Patrick Pérez; Gaël Richard*
- Looking to Listen at the Cocktail Party: A Speaker-Independent Audio-Visual Model for Speech Separation, *Tali Dekel; Miki Rubinstein; Inbar Mosseri; William T. Freeman; Oran Lang; Kevin Wilson; Ariel Ephrat; Avinatan Hasidim*
- The Excitement of Sports: Automatic Highlights Using Audio/Visual Cues, *Michele Merler; Dhiraj Joshi; Khoi-Nguyen C. Mac; Quoc-Bao Nguyen; Stephen Hammer; John Kent; Jinjun Xiong; Minh N. Do; John R. Smith; Rogerio S. Feris*
- A Multimodal Approach to Mapping Soundscapes, *Tawfiq Salem; Menghua Zhai; Scott Workman; Nathan Jacobs*
- Multimodal Attention for Fusion of Audio and Spatio-temporal Features for Video Description, *Chiori Hori; Takaaki Hori; Gordon Wichern; Jue Wang; Teng-Yok Lee; Anoop Cherian; Tim K. Marks*
- Visual Rhythm and Beat, *Abe Davis; Maneesh Agrawala*
- Inverting Audio-Visual Simulation for Shape and Material Perception, *Zhoutong Zhang; Jiajun Wu; Qiujia Li; Zhengjia Huang; Joshua B. Tenenbaum; William T. Freeman*

1600 **Invited Talk:** *William Freeman (MIT/Google)*

1630 **Invited Talk:** *Relja Arandjelovic (DeepMind)*

1700 Paper Session 3

Notes:

Panel: How to be a Good Citizen of the CVPR Community

Organizers: Devi Parikh
Dhruv Batra

Location: Ballroom E

Schedule: Full Day

0900 **Opening Remarks**

0910 **Invited Talk:** How to Write a Good CVPR Submission, *Bill Freeman (Google & MIT)*

0930 **Invited Talk:** Rights and Obligations, *Katsushi Ikeuchi (Microsoft Research & Univ. of Tokyo)*

0950 **Invited Talk:** Creating an Inclusive and Welcoming Culture at CVPR, *Timnit Gebru (Microsoft Research)*

1010 **Invited Talk:** Strengthening our Community Through Mentorship, Leadership, and Inclusiveness, *Sven Dickinson (Univ. of Toronto)*

1030 Morning Break (Halls 1-4)

1100 **Invited Talk:** TBA, *Kristen Grauman (Univ. of Texas at Austin)*

1120 **Invited Talk:** Doing Good Research. Or Reviewing., *Vladlen Koltun (Intel)*

1140 **Invited Talk:** Research in Context, Impact of Persuasive Images on Society, Education and Outreach, Research With Undergrads, Fostering and Mentoring Students, *Adriana Kovashka (Univ. of Pittsburgh)*

1200 Lunch (Halls 1-4)

1330 **Invited Talk:** How to Write a Good Paper, *Jitendra Malik (Facebook AI Research & Univ. of California at Berkeley)*

1350 **Invited Talk:** How to Be a Good Reviewer / Area Chair, How to Do Good Research, Proper Evaluation, *Cordelia Schmid (INRIA & Google)*

1410 **Invited Talk:** How to Win Colleagues and Influence People, *Derek Hoiem (Reconstruct & Univ. of Illinois at Urbana-Champaign)*

1430 **Invited Talk:** The Importance of Collaborations Between Researchers, Reproducibility of Published Work (Including Open Sourcing), *Georgia Gkioxari (Facebook AI Research)*

1450 **Invited Talk:** Welcome 大家, *Michael Brown (York Univ.)*

1510 **Invited Talk:** What PCs Told ACs for CVPR 18, *David Forsyth (Univ. of Illinois at Urbana-Champaign)*

1530 Afternoon Break (Halls 1-4)

1600 **Panel:** How to Be a Good Citizen of the CVPR Community, *Panelists: Timnit Gebru (Microsoft Research), David Forsyth (Univ. of Illinois at Urbana-Champaign), (Tentative) Kristen Grauman (Univ. of Texas at Austin), Michael Brown (York Univ.), Derek Hoiem (Reconstruct & Univ. of Illinois at Urbana-Champaign), Georgia Gkioxari (Facebook AI Research), Katsushi Ikeuchi (Microsoft Research & Univ. of Tokyo), Jitendra Malik (Facebook AI Research & Univ. of California at Berkeley), and Cordelia Schmid (INRIA & Google)*

1725 **Closing Remarks**

Notes:

GOLD SPONSORS

SILVER SPONSORS

DIAMOND SPONSORS

amazon

Baidu 百度

Face++ 旷视

facebook

Google

Microsoft

NVIDIA

商汤 sensetime

PLATINUM SPONSORS

Alibaba Group 阿里巴巴集团

蚂蚁金服 ANT FINANCIAL

AURORA

ByteDance 字节跳动

TOUTIAO AI LAB 今日头条人工智能实验室

COGNEX

DiDi

HUAWEI

IBM

intel AI

lyft

MOMENTA

NAVER | LINE

oppo

QUALCOMM

SAMSUNG

Tencent 腾讯

UBER

XILINX ALL PROGRAMMABLE

依图 | YITU