ECCV 2018

European Conference on Computer Vision

8 – 14 September 2018 | Munich, Germany

Program

Contents

Welcome note / Foreword	3
ECCV Proceedings	4
General Information / Organizer	7
Time Schedule Gasteig	9
Time Schedule Workshops & Tutorials	10
About Munich	12
Venue Information	14
City Map	16
Presenter Instructions	18
Restaurant Suggestions	20
Floor Plan GASTEIG Cultural Center	22
Floor Plan TU Munich	25
Workshops /Tutorials – Saturday, 8 September 2018	26
Workshops /Tutorials – Sunday, 9 September 2018	56
Program Main Conference	84
Demo Sessions	96
Workshops / Tutorials – Friday, 14 September 2018	122
Poster Sessions	141
Sponsors and Exhibitors	247
Exhibition Plan	251
Evening Program	252

It is our great pleasure to host the European Conference on Computer Vision 2018 in Munich, Germany. This constitutes by far the largest ECCV ever. With near 3200 registered participants and another 650 on the waiting list as we write, participation has more than doubled since the last ECCV in Amsterdam. We believe that this is due to a dramatic growth of the computer vision community combined with the popularity of Munich as a major European hub of culture, science and industry. The conference takes place in the heart of Munich in the concert hall Gasteig with workshops and tutorials held in the downtown campus of the Technical University of Munich.

One of the major innovations for ECCV18 is the free perpetual availability of all conference and workshop papers, which is often referred to as open access.[1] Since 2013, CVPR and ICCV have had their papers hosted by the Computer Vision Foundation (CVF), in parallel with the IEEE Xplore version. This has proved highly beneficial to the computer vision community.

We are delighted to announce that for ECCV18 a very similar arrangement has been put in place with the cooperation of Springer. In particular, the author's final version will be freely available in perpetuity on a CVF page, while SpringerLink will continue to host a version with further improvements, such as activating reference links and including video. We believe that this will give readers the best of both worlds; researchers who are focused on the technical content will have a freely available version in an easily accessible place, while subscribers to SpringerLink will continue to have the additional benefits that this provides. We thank Alfred Hofmann from Springer for helping to negotiate this agreement, which we expect will continue for future versions of ECCV.

Horst Bischof
General Chair

Daniel Cremers
General Chair

Bernt Schiele
General Chair

Ramin Zabih
General Chair

1 This is not precisely the same use of the term as in the Budapest declaration

Welcome to the proceedings of the 2018 European Conference on Computer Vision (ECCV 2018) held in Munich, Germany. We are delighted to present this volume reflecting a strong and exciting program, the result of anextensive review process. In total, we received 2439 valid paper submissions. 776 were accepted (31.8 %): 717 as posters (29.4%), and 59 as oral presentations (2.4 %). All orals are presented as posters as well. The program selection process was complicated this year by the large increase in the number of submitted papers, +65% over ECCV 2016, and the use of CMT3 for the first time for a computer vision conference. The program selection process was supported by four program co-chairs (PCs), 126 area chairs (ACs), and 1199 reviewers with reviews assigned.

We were primarily responsible for the design and execution of the review process. Beyond administrative rejections, we were involved in acceptance decisions only in the very few cases where the ACs were not able to agree on a decision. As PCs, as is customary in the field, we were not allowed to co-author a submission. General co-chairs and other co-organizers who played no role in the review process, were permitted to submit papers, and were treated as any other author.

Acceptance decisions were made by two independent ACs. The ACs also made a joint recommendation for promoting papers to Oral status. We decided on the final selection of Orals based on the ACs' recommendations. There were 126 ACs, selected according to their technical expertise, experience, and geographical diversity (63 from European, 9 from Asian/ Australian, and 54 from North American institutions). 126 ACs is a substantial increase in the number of ACs due to the natural increase in the number of papers and to our desire to maintain the number of papers assigned to each AC to a manageable number to ensure quality. The ACs were aided by the 1199 reviewers to whom papers were assigned for reviewing. The Program Committee was selected from committees of previous ECCV, ICCV, and CVPR conferences and was extended on the basis of suggestions from the ACs. Having a large pool of Program Committee members for reviewing allowed us to match expertise while reducing reviewer loads. No more than eight papers were assigned to a reviewer, maintaining the reviewers' load at the same level as ECCV 2016 despite the increase in the number of submitted papers.

Conflicts of interest between ACs, Program Committee members, and papers were identified based on the home institutions, and on previous collaborations of all researchers involved. To find institutional conflicts, all authors, Program Committee members, and ACs were asked to list the Internet domains of their current institutions. We assigned on average approximately 18 papers to each AC. The papers were assigned using the affinity scores from the Toronto Paper Matching System (TPMS) and additional data from the OpenReview system, managed by a UMass group. OpenReview used additional information from ACs' and authors' records to identify collaborations and to generate matches. OpenReview was invaluable in refining conflict definitions and in generating quality matches. The only glitch is that, once the matches were generated, a small percentage of

papers were unassigned because of discrepancies between the OpenReview conflicts and the conflicts entered in CMT3. We manually assigned these papers. This glitch is revealing of the challenge of using multiple systems at once (CMT3 and OpenReview in this case), which needs to be addressed moving forward.

After assignment of papers to ACs, the ACs suggested seven reviewers per paper from the Program Committee pool. The selection and rank ordering was facilitated by the TPMS affinity scores visible to the ACs for each paper/reviewer pair. The final assignment of papers to reviewers was generated again through OpenReview in order to account for refined conflict definitions. This required new features in the OpenReview matching system toaccommodate the ECCV workflow, in particular to incorporate selection ranking, and maximum reviewer load. Very few papers received fewer than three reviewers after matching and were handled through manual assignment. Reviewers were then asked to comment on the merit of each paper and to make an initial recommendation ranging from definitely reject to definitely accept, including a borderline rating. The reviewers were also asked to suggest explicit questions they wanted to see answered in the authors' rebuttal. The initial review period was 5 weeks. Because of the delay in getting all the reviews in, we had to delay the final release of the reviews by four days. However, because of the slack built-in at the tail end of the schedule, we were able to maintain the decision target date with sufficient time for all the phases. We reassigned over 100 reviews from 40 reviewers during the review period. Unfortunately, the main reason for these reassignments was reviewers declining to review, after having accepted to do so. Other reasons included technical relevance and occasional unidentified conflicts. We express our thanks to the emergency reviewers who generously accepted to perform these reviews under short notice. In addition, a substantial number of manual corrections had to do with reviewers using a different email address than the one that was used at the time of the reviewer invitation. This is revealing of a broader issue with identifying users by email addresses which change frequently enough to cause significant problems during the timespan of the conference process.

The authors were then given the opportunity to rebut the reviews, to identify factual errors and to address the specific questions raised by the reviewers over a seven day rebuttal period. The exact format of the rebuttal was the object of considerable debate among the organizers, as well as with prior organizers. At issue is to balance giving the author the opportunity to respond completely and precisely to the reviewers, e.g., by including graphs of experiments, while avoiding requests for completely new material, experimental results, not included in the original paper. In the end, we decided on the two-page PDF document in conference format. Following this rebuttal period, reviewers and ACs discussed papers at length, after which reviewers finalized their evaluation and gave a final recommendation to the ACs. A significant percentage of the reviewers did enter their final recommendation if it did not differ from their initial recommendation. Given the tight schedule, we did not wait until all were entered.

After this discussion period, each paper was assigned to a second AC. The AC/paper matching was again run through OpenReview. Again, the OpenReview team worked quickly to implement the features specific to this process, in this case accounting for the existing AC assignment, as well as minimizing the fragmentation across ACs, so that each AC had on average only 5.5 buddy ACs to communicate with. The largest number was 11. Given the complexity of the conflicts, this was a very efficient set of assignments from OpenReview. Each paper was then evaluated by its assigned pair of ACs. For each paper, we required each of the two ACs assigned to certify both the final recommendation and the metareview (aka consolidation report.) In all cases, after extensive discussions, the two ACs arrived at a common acceptance decision. We maintained these decisions, with the caveat that we did evaluate, sometimes going back to the ACs, a few papers, for which the final acceptance decision substantially deviated from the consensus from the reviewers, amending three decisions in the process.

We want to thank everyone involved in making ECCV 2018 possible. The success of ECCV 2018 depended on the quality of papers submitted by the authors, and on the very hard work of the ACs and the Program Committee members. We are particularly grateful to the OpenReview team (Melisa Bok, Ari Kobren, Andrew McCallum, Michael Spector) for their support, in particular their willingness to implement new features, often on a tight schedule, to Laurent Charlin for the use of the Toronto Paper Matching System, to the CMT3 team, in particular in dealing with all the issues that arise when using a new system, to Friedrich Fraundorfer and Quirin Lohr for maintaining the online version of the program, and to the CMU staff (Keyla Cook, Lynnetta Miller, Ashley Song, Nora Kazour) for assisting with data entry/editing in CMT3. Finally, the preparation of these proceedings would not have been possible without the diligent effort of the publication chairs, Albert Ali Salah and Hamdi Dibeklioglu, and of Anna Kramer and Alfred Hofman from Springer.

General Chairs

Horst Bischof	Graz University of Technology, Austria
Daniel Cremers	Technical University of Munich, Germany
Bernt Schiele	Saarland University, Max Planck Institute for Informatics, Germany
Ramin Zabih	CornellNYCTech, USA

Program Chairs

Vittorio Ferrari	Google Research and University of Edinburgh, UK
Martial Hebert	Carnegie Mellon University,USA
Cristian Sminchisescu	Lund University, Sweden
Yair Weiss	Hebrew University, Israel

Congress Organisation

Interplan Congress, Meeting & Event Management AG
Project Management: Jana Bylitza
Landsberger Straße 155, 80687 Munich
Phone +49 (0) 89 54 82 34 62
eccv2018@interplan.de

Opening Hours:

Registration Desk	08. September 2018	07.00 am – 06.30 pm
	09. September 2018	07.30 am – 06.30 pm
	10. September 2018	07.30 am – 07.00 pm
	11. September 2018	07.30 am – 06.30 pm
	12. September 2018	08.00 am – 06.30 pm
	13. September 2018	08.00 am – 06.30 pm
	14. September 2018	08.00 am – 06.30 pm
Industry exhibition	10. September 2018	08.30 am – 06.00 pm
	11. September 2018	08.30 am – 06.00 pm
	12. September 2018	08.30 am – 06.00 pm
	13. September 2018	08.30 am – 06.00 pm

Cloak rooms/luggage deposit

You may bring backpacks and handbags to the GASTEIG with you. Larger items such as suitcases will not be allowed in the lecture halls. These items need to be checked at the wardrobe, which will be located in the basement. The wardrobe will be supervised during the opening hours of the congress.

Please note that there is no wardrobe at the TU Munich. If possible, please leave your suitcase at the hotel and do not bring it to the university with you. If you have to bring a suitcase to the venue, you have to take it to the workshop rooms with you. We cannot watch it for you.

Lost and found

There is a lost and found service at the registration desk in both venues.

Questions and requests

Please feel free to ask the Staff – wearing branded Staff ECCV-T-Shirts – if you have any question or request. They will provide assistance to speakers and other participants with practical answers.

Wifi

Free WiFi will be available on site for all ECCV 2018 attendees. This is the log in information. To ensure that all participants can benefit from the WiFi connection, we kindly ask you not to stream or download movies etc.

Gasteig Cultural Center
Network: ECCV_2018
Password: ECCV2018_Munich

TU Munich
Preferably, just use the EDUROAM network if you have eduroam access. Otherwise, you can also log in to the free „BayernWLAN“ network.

MO	TUE	WED	THU
08:30am - 09:45am Orals 1	08:30am - 09:45am Orals 1	08:30am - 09:45am Orals 1	08:30am - 09:45am Orals 1
09:45am - 10:00am Break	09:45am - 10:00am Break	09:45am - 10:00am Break	09:45am - 10:00am Break
10:00am - 12:00pm Poster 1	10:00am - 12:00pm Poster 1	10:00am - 12:00pm Poster 1	10:00am - 12:00pm Poster 1
12:00pm - 01:00pm Lunch	12:00pm - 01:00pm Lunch	12:00pm - 01:00pm Lunch	12:00pm - 01:00pm Lunch
01:00pm - 02:15pm Orals 2	1:00pm - 02:15pm Orals 2	1:00pm - 02:15pm Orals 2	01:00pm - 02:15pm Orals 2
02:15pm - 02:45pm Break	02:15pm - 02:45pm Break	02:15pm - 02:30pm Break	02:15pm - 02:45pm Break
02:45pm - 04:00pm Orals 3	02:45pm - 04:00pm Orals 3	02:30pm - 04:00pm Poster 2 04:00pm - 05: 15pm Orals 3	02:45pm - 04:00pm Orals 3
04:00pm - 06:00pm Poster 2	04:00pm - 06:00pm Poster 2	05:15pm - 6:45pm Poster 3	04:00pm - 06:00pm Poster 2

	Saturday 08 morning	Saturday 8 afternoon
Audimax 0980	Workshop on Shortcomings in Vision and Language	
1200 Carl von Linde Hörsaal	3D sensing and Reconstruction	
2750 Karl Marx von Bauernfeind		
N1179	Adversarial Machine Learning	Instance-level Visual Recognition
N1189	Normalization Methods for Training Deep Neural Networks: Theory and Practice	Brain-Driven Computer Vision
N1070ZG	Visual Localization: Feature-based vs. Learned Approaches	PoseTrack Challenge: Articulated People Tracking in the Wild
N1080ZG	Representation Learning for Pedestrian Re-identification	Anticipating Human Behavior“
Theresianum 602	Face Tracking an its Applications	Open Images Challenge Workshop
Theresianum 606	HoloLens	AutoNUE: Autonomous Navigation in Unconstrained Environments
N1090ZG	Human Identification at a Distance by Gait and Face Analysis	WIDER Face and Pedestrian Challenge
N1095ZG	Vision Meets Drone: A Challenge	UAVision
Theresianum 601 0506. EG.601	Video Recognition and Retrieval at the TRECVID Benchmark	Vision for XR
Theresianum 1601	Functional Maps: A Flexible Representation for Learning and Computing Correspondence	Workshop on Objectionable Content and Misinformation

Sunday 09 morning	Sunday 09 afternoon	Friday 14 morning	Friday 14 afternoon
Workshop on Human Behavior Understanding (HBU) - Focus Theme: Towards Generating Realistic Visual Data of Human Behavior	360-degree Perception and Interaction Workshop		
ApolloScape: 3D Understanding for Autonomous Driving		Third International Workshop on Video Segmentation	1st Large-scale Video Object Segmentation Challenge
		3rd Workshop on Geometry Meets Deep Learning	
3D Reconstruction Meets Semantics	6th Workshop Assistive Computer Vision and Robotics	The Visual Object Tracking Challenge Workshop VOT2018	PeopleCap 2018: Capturing and Modeling Human Bodies, Faces and Hands
2nd Compact and Efficient Feature Representation and Learning in Computer Vision		6th Computer Vision for Road Scene Understanding and Autonomous Driving	
2nd YouTube-8M Large-Scale Video Understanding Challenge Workshop		Workshop and Challenge on Perceptual Image Restoration and Manipulation	Bias Estimation in Face Analytics (BEFA)
Joint COCO and Mapilary Recognition Challenge Workshop		3D Reconstruction in the Wild	
Multimodal Learning and Applications Workshop	1st Person in Context (PIC) Workshop and Challenge	First Workshop on Computer Vision For Fashion, Art and Design	
4th International Workshop on Recovering 6D Object Pose	4th International Workshop on Observing and Understanding Hands in Action	What is Optical Flow for?	
Egocentric Perception, Interaction and Computing (EPIC)	Women in Computer Vision Workshop	1st Workshop on Interactive and Adaptive Learning in an Open World	
Workshop on Visual Learning and Embodied Agents in Simulation Environments		5th TASK-CV: Transferring and Adapting Source Knowledge in Computer Vision and 2nd VisDA Challenge	
Perceptual Organization in Computer Vision (POCV)		VizWiz Grand Challenge: Answering Visual Questions from Blind People	
4th Workshop on Computer VISion for ART Analysis (VISART IV)		BioImage Computing	

Discover Munich by attending the ECCV 2018. This beautiful city offers an extraordinary mixture of international cultural excellence with its museums and concert halls, as well as true Bavarian tradition and hospitality all set in the breathtaking pre-alpine landscape of lakes, mountains and fairytale castles.

The townscape – Inviting, historical and visionary

King Ludwig I influenced Munichs townscape with his pompous and generous architecture with broad avenues and the contrast between classicistic restraint and baroque profusion. Bold creativity and innovation have set new architectural accents all over the city. The world-renowned tent roof, the landmark of the Olympic Park, inspires spectators even four decades after its construction. The Allianz Arena is regarded Germany's most beautiful and thrilling soccer stadium. The recently opened BMW Welt is a milestone of dynamic architecture.

The downtown area presents with a clear, delightful and charming scenery marked by the distinctive feature of the Frauenkirche – church with it's towers rising high above the city roofs.

Live and let live

Munich's appeal and success are closely linked to the joy of life, cosiness and a cosmopolitan atmosphere. This is THE place for gourmets, night revelers and party animals.

„Eating and drinking keeps body and soul together". In Munich this old proverb still holds good today. You will find cosy Bavarian inns, international cuisine, gourmet restaurants next to our world-famous beer gardens, where people of all age, families, singles, couples and guests from all over the world get together under shady chestnut trees and enjoy a stein of freshly tapped beer.

Shopping in Munich is a particular pleasure. Strolling along the extravagant Maximilianstrasse and Theatinerstrasse you quickly reach the pedestrian mall where department stores and shops offer their selected goods. Trendy and fancy merchandise is found, for example, in the Gärtnerplatz and Glockenbach quarters; a variety of souvenirs is available at the „Platzl" near the Hofbräuhaus; culinary treats from all over the world are best bought downtown at Viktualienmarkt. The famous Munich nightlife entices with numerous pubs, bars and clubs.

Facts about Munich...

- was founded in 1158
- it is the third largest city in Germany after Berlin and Hamburg and has a population of about 1,4 million inhabitants
- offers more than 50 museums to the interested visitor as well as over 60 theatres and cabarets
- presents the “Neues Rathaus” which is one of Germany´s most distinctive buildings and which is home to the famous Munich glockenspiel which chimes at 11am, 12pm and 5pm
- is home to more than 50.000 students at 13 different universities
- enjoys an excellent international reputation as a competence centre in research, science and medicine
- is a green city. The nature within the city offers
  - surfing downtown on the river Isar
  - rafting from beergarden to beergarden
  - getting a tan in the „English Garden“ – which is bigger than the NY Central Park
  - jumping into its rivers and various lakes

And, last but not least:
- Munich is City of the Oktoberfest. The largest beer festival in the World – with more than 7 million visitors every year (starting from 22 September 2018)

The ECCV 2018 conference will take place at the locations described below. A map showing all locations is provided.

A Technische Universität München (TU)

For **Workshops and Tutorials (8, 9 and 14 September)**
Conference Registration

Address Arcisstraße 21, 80333 Munich

Direction To get to the TU München please take any S-Bahn to the Munich Central Station. From the Central Station please use the subway “U2” (direction: Feldmoching) and get off at the second station “Theresienstraße”.

To get to the TU München please follow the street “Theresienstraße”. (East, where the street numbers went down) and turn right at the third street “Arcisstraße”. On the right side, you will find the entrance of the TU München.

B GASTEIG Cultural Center

For **Main conference (10, 11, 12 and 13 September)**
Conference Registration
Welcome Reception (10 September)

Address Rosenheimer Str. 5, 81667 Munich

Direction To get to the Gasteig please take any S-Bahn to the “Rosenheimer Platz”. Otherwise, you can take the Tramline 17 to the station „Am Gasteig“.

At the “Rosenheimer Platz” please follow the »Gasteig« signage on the platform and in the station.

C

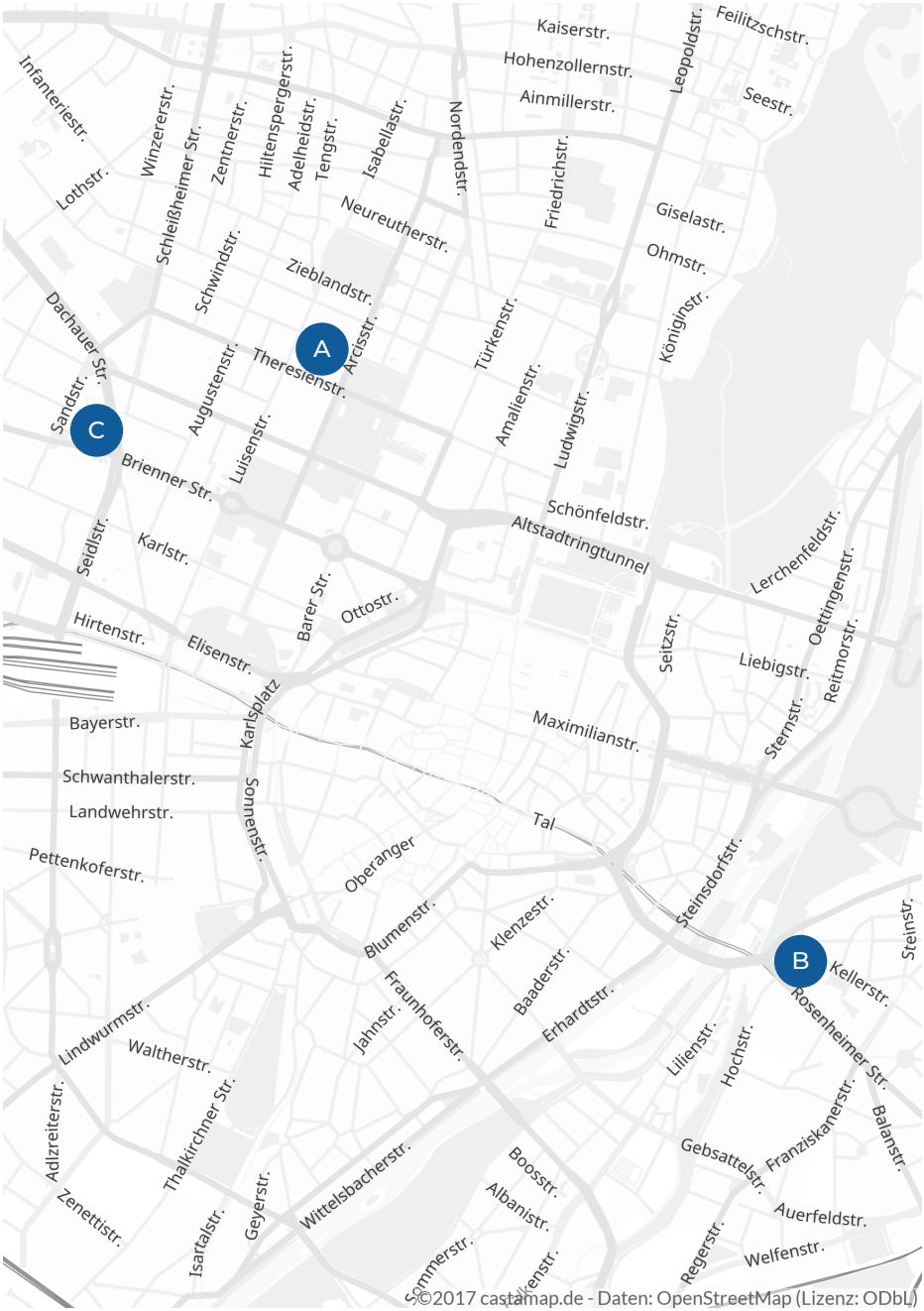
Löwenbräukeller München

For **Congress Dinner (12 September 2018)**

Address Nymphenburger Straße 2, 80335 Munich

Direction To get to the Löwenbräukeller please take any S-Bahn to the Munich Central Station. From the Central Station please use the subway “U1” (direction: Olympia-Einkaufszentrum) and get off at the first station “Stiglmaierplatz”.

The stop “Stiglmaierplatz” is directly at the Löwenbräukeller.



Orals:

Time:
Each paper in an oral session is allocated 13 minutes. Additional 2 minutes are allocated for questions. You have to leave the podium once your time is up. Do not exceed the given time limit.

Your presentation must be submitted as Windows or Mac compatible PowerPoint files on PC-readable CDs, DVDs, external disk drives, USB sticks or memory sticks. Presentation format is 16:9.
Microsoft ppt or keynote ppt are accepted. In case you want to present with another format, please inform the media check about this. Do not go directly to the stage to present with your own laptop as the technicians at the media check have to be informed before.

All oral presentations have also been allocated a poster presentation. The poster presentation will be scheduled in the subsequent poster session after the talk.

Poster:

Please note that we have two poster areas. Make sure that you hang up your poster in the correct poster area. Volunteers at the poster desk in each poster area will help you to find the correct poster board. They will furthermore hand out tape in case you need something.

Poster numbers 1 – 53 are allocated on the second floor of the Philharmonie foyer.
Poster numbers 54 – 91 are allocated on the first floor of the Carl Orff foyer.

Presentation format:
The poster format is A0 landscape. Poster boards are 1940mm wide and 950mm tall. Adhesive material and/or pins will be provided for mounting the posters to the boards. If you have special requirements, please contact poster chairs as soon as possible. We will try to accommodate your requests as much as possible.

Arrival and take off time
Poster presenters are asked to install their posters during the coffee break prior to the poster session. Posters are to be removed from the poster boards by the presenters at the end of the session. If not removed, volunteers will collect the remaining poster.

Poster Session Guidelines (Workshops/Tutorials)

Poster papers have been assigned to poster boards as indicated in the program. Please note that we have two poster areas in two rooms on the second floor of the Theresianum (room 2605 and 2607). A volunteer in front of the rooms can help you to find the correct poster board and will hand out tape in case you need something.

Presentation format

Poster boards are 95cm wide and 1.84m tall. Posters with the format PORTRAIT 1.00m (width) x 1.20m (height) would be possible. Much easier would be the format 95cm x 1.84m. Adhesive material and/or pins will be provided for mounting the posters to the boards.
In case you want to print your poster in Munich directly, please find below a possibility of copy shop in Munich: https://www.printy.de/en/

Arrival and take off time

Poster presenters are asked to install their posters before the poster session starts and to remove it from the poster boards at the end of the session.

Demo Presentation Guidelines (Main Conference)

The demos will be presented in Chorprobensaal at one of the designated demo stations #1 to #4. Make sure that you set up your demo at the correct demo station. Volunteers in Chorprobensaal will help you to find the correct demo station. They will furthermore hand out tape for the posters in case you need something.

Presentation Format

The conference will provide the following basic equipment at each demo station: 1 monitor with HDMI plug, 1 table, 2 chairs, 2 power outlets, 1 poster board. Space for the demo is limited to a 1,5m x 2m area at the demo station. The poster format is DIN A0 landscape. Poster boards are 1940mm wide and 950mm tall. Adhesive material and/or pins will be provided for mounting the posters to the boards.
Please make sure that the visitors can actually experience a live demo of your computer vision system (and not just a video playing on a monitor or a poster presentation).

Arrival and Take Off Time

Demo presenters are asked to setup their demo and install their posters during the coffee break prior to the demo session. Additional demo equipment brought by the presenters and the posters are to be removed from the demo station by the presenters at the end of the session. If not removed, volunteers will collect the remaining items.

Oral Session Guidelines (Workshops/Tutorials)

Presenters are asked bring their own laptops to present their slides. Please make sure that you are on time inside the room. The TU offers VGA and HDMI plugs.

Restaurant Suggestions close to the TU Munich

Catering TU Munich Foyer AUDIMAX
Arcisstraße 21, 80333 Munich
Open daily: 08am – 06pm
Coffeestation
Self-paying Lunch

Tenmaya
Theresienstraße 43, 80333 Munich
Open daily: 11.30am – 03pm
05.30pm – 11.30pm
Distance: 2 min. by foot
Japanese restaurant

Steinheil 16
Steinheilstraße 16, 80333 Munich
Open daily: 10am – 01am
Distance: 5 min. by foot
Traditional Bavarian food

Kims Restaurant
Theresienstraße 138, 80333 Munich
Open daily (except Mondays):
11.30am – 02pm & 06pm – 11pm
Distance: 6 min. by foot
Korean restaurant

Theresa Grill
Theresienstraße 29, 80333 München
Open daily: 6pm – 01am
Distance: 8 min. by foot
Steakhouse

The Italian Shot
Theresienstraße 40, 80333 Munich
Open daily: 12pm – 03pm &
05pm – 01am
Distance: 8 min. by foot
Italian restaurant

Hamburgerei EINS
Brienner Str. 49, 80333 Munich
Open daily: 11.30am – 10pm
Distance: 9 min. by foot
Burger

Türkenhof
Türkenstraße 78, 80799 Munich
Open daily: 11am – 01am
Distance: 11 min. by foot
Traditional Bavarian food

Restaurant Suggestions close to the GASTEIG Cultural Center

gast
Rosenheimer Straße 5,
81667 Munich
Open daily: 11am – 01am
Distance: 1 min. by foot
Italian and Asian Restaurant

Kuchlverzeichnis
Rosenheimer Straße 10,
81667 Munich
Open daily: 05.30am – 01am
Distance: 3 min. by foot
Traditional Bavarian food

Rosi Kaffeehaus & Bar
Rosenheimer Straße 2,
81669 Munich
Open daily: 08am – 01am
Distance: 4 min. by foot
Snack Bar

Kam Yi
Rosenheimer Straße 32, 81669 Munich
Open daily: 11.30am – 11pm
Distance: 2 min. by foot
Chinese Restaurant

Chez Fritz
Preysingstraße 20, 81667 Munich
Open daily: 05.30pm – 01am
Distance: 6 min. by foot
French Restaurant

Lollo Rosso Bar(varain) Grill
Milchstraße 1, 81667 Munich
Open daily: 05pm – 01am
Distance: 5 min. by foot
Traditional Bavarian food

L´Osteria
Innere Wiener Straße 2, 81667 München
Open daily: 11.30am – 11pm
Distance: 4 min. by foot
Italian Restaurant

Hofbäukeller
Innere Wiener Straße 19, 81667 Munich
Open daily: 10am – 12pm
Distance: 8 min. by foot
Traditional Bavarian food

Wirtshaus in der Au
Lilienstraße 5, 81667 München
Open daily: 05pm – 12pm
Distance: 7 min. by foot
Traditional Bavarian food

Restaurant Del Cavaliere
Weißenburger Str. 3, 81667 München
Open daily: 11.30am – 00.30am
Distance: 5 min. by foot
Italian Restaurant

Vegetarian & Vegan Restaurants

Emmi´s Kitchen
Rosenheimer Str. 67, 81667 Munich
Open daily: 11am – 9pm

Max Pett
Pettenkoferstraße 8, 80336 Munich
Open daily: 11.30am – 11pm

Prinz Myshkin
Hackenstraße 2, 80331 Munich
Open daily: 12 – 10pm

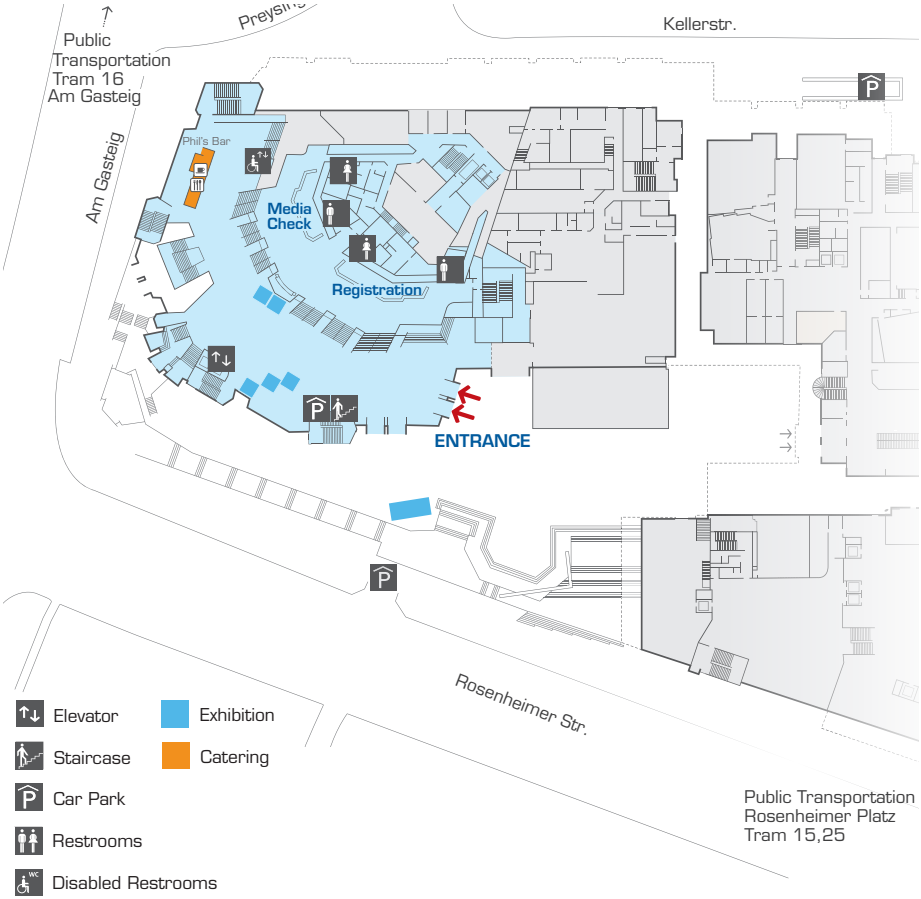
Halal Restaurants

Myra Restaurant
Thalkirchner Str. 145, 81371 Munich
Open daily: 05pm – 01am

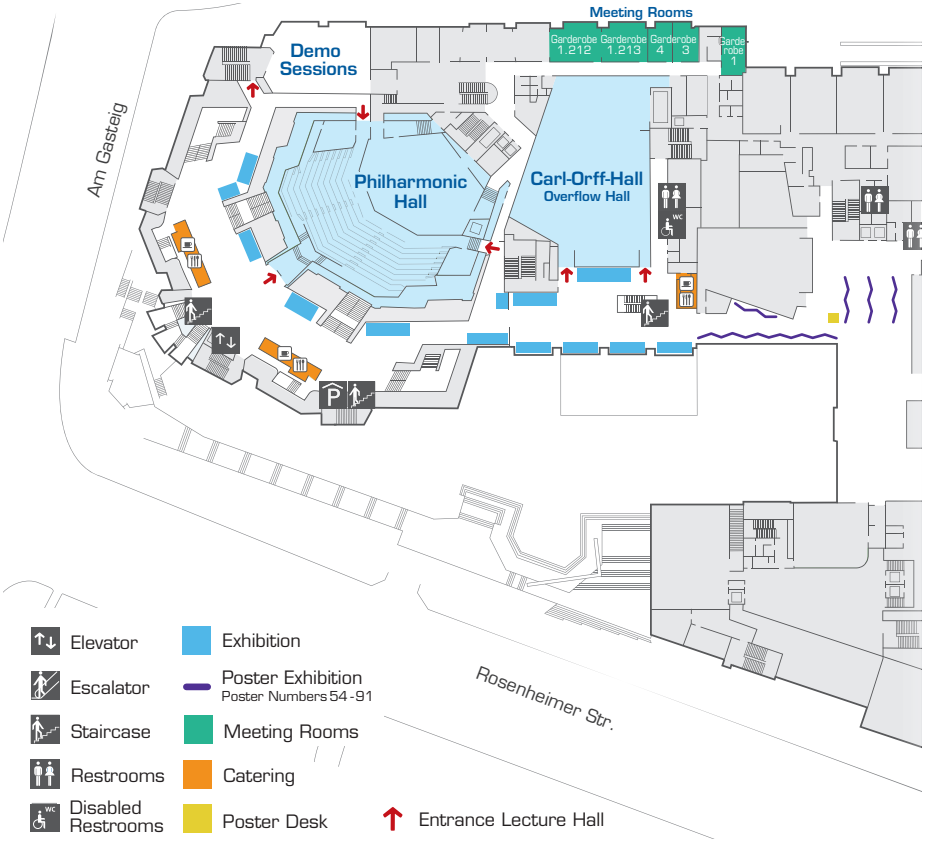
Pardi
Volkartstraße 24, 80634 München
Open daily: 09 - 01am

Arabesk
Kaulbachstraße 86, 80802 München
Open daily: 06 – 11pm

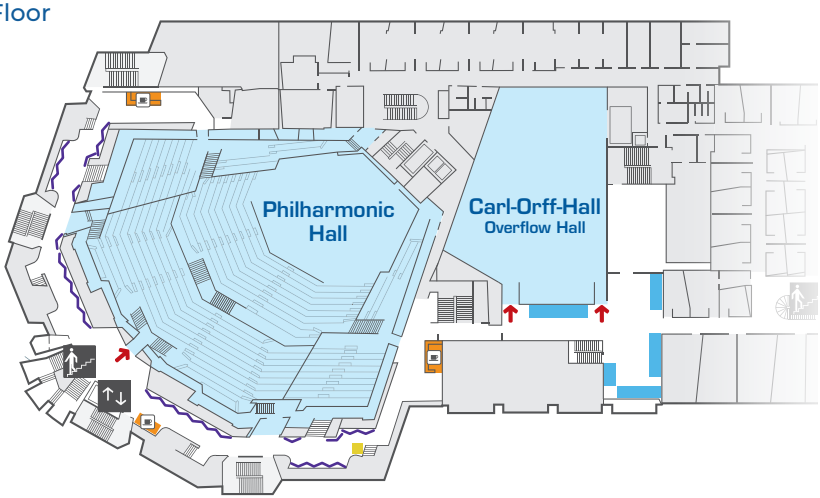
Ground Floor



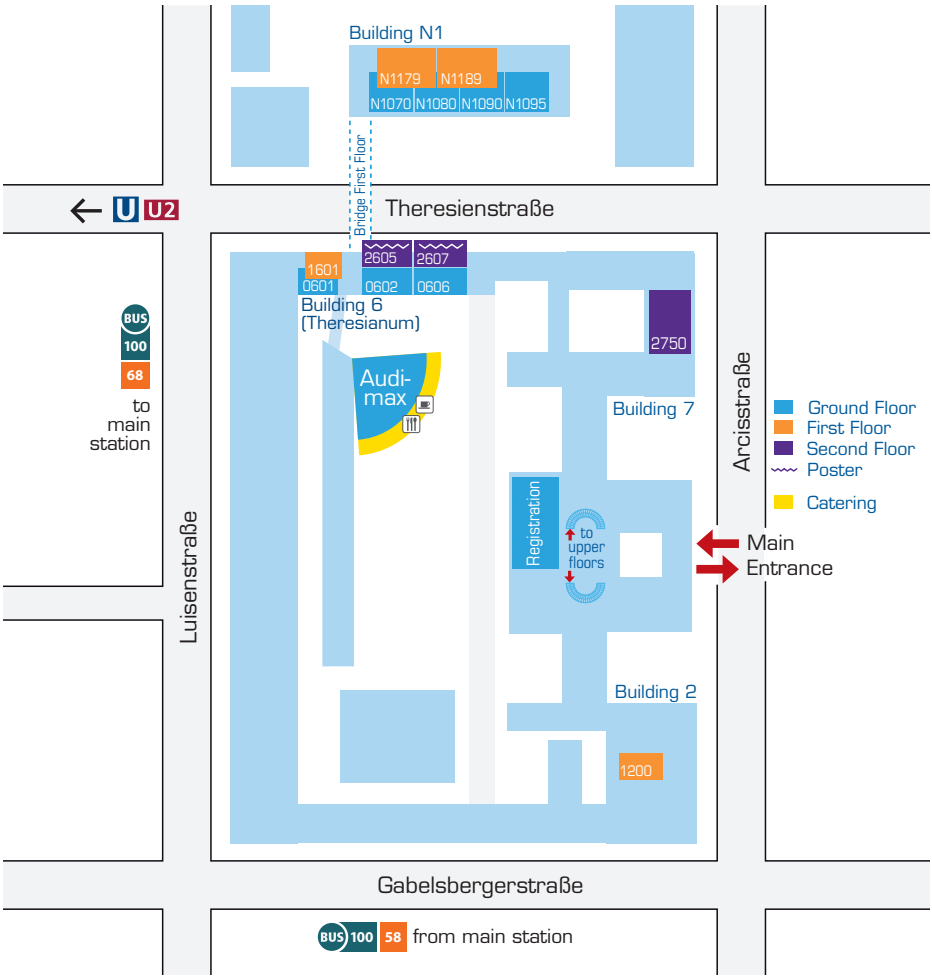
First Floor


Meeting Rooms
Garderobe 1.212
Garderobe 1.213
Garderobe 4
Garderobe 3
Garderobe 1
Demo Sessions
Philharmonic Hall
Carl-Orff-Hall
Overflow Hall
Am Gasteig
Rosenheimer Str.
Elevator
Escalator
Staircase
Restrooms
Disabled Restrooms
Exhibition
Poster Exhibition
Poster Numbers 54 - 91
Meeting Rooms
Catering
Poster Desk
Entrance Lecture Hall

Second Floor



Elevator
Staircase
Exhibition
Poster Exhibition
Poster Numbers 1-53
Catering
Poster Desk
Entrance Lecture Hall


Building N1
N1179
N1189
N1070
N1080
N1090
N1095
Bridge First Floor
Theresienstraße
U U2
BUS 100 68
to main station
1601
0601
2605
2607
0602
0606
Building 6 (Theresianum)
Audi-max
2750
Building 7
Arcisstraße
Ground Floor
First Floor
Second Floor
Poster
Catering
Registration
to upper floors
Main Entrance
Luisenstraße
Building 2
1200
Gabelsbergerstraße
BUS 100 58 from main station

Tutorial Number: T1

Human Identification at a Distance by Gait and Face Analysis

Date: Saturday 8th morning

Room: N1090ZG

Organizers: Yongzhen Huang, Liang Wang, Man Zhang, Tieniu Tan

SCHEDULE

09:00 – 09:20 Overview of the tutorial: Motivations, challenges, datasets and evaluation.

09:20 – 10:30 Gait and face analysis for human identification at a distance.

1. A brief review of face analysis.
2. Traditional approaches of gait analysis.
   a) Feature representation and classification.
   b) Comparison of different methods.
3. Deep learning-based approaches for gait analysis.
   a) CNN for gait recognition
   b) GAN for gait recognition
   c) Cross-view gait recognition
   d) Soft biometrics in gait recognition

10:30 – 11:00 Coffee Break

11:00 – 11:20 Applications of gait and face analysis at a distance.

11:20 – 11:40 Experience our newest human identification system.

11:40 – 12:00 Conclusion, open questions and discussion.

Tutorial Number: T2

Video Recognition and Retrieval at the TRECVID Benchmark

Date: Saturday 8th morning

Room: Theresianum 601 EG

Organizers: George Awad, Alan Smeaton, Cees Snoek, Shin'ichi Satoh, Kazuya Ueki

SCHEDULE

08:00	Tutorial starts
08:00 – 08:45	Introduction to TRECVID (George Awad)
08:45 – 09:30	Video To Text (Alan Smeaton)
09:30 – 10:15	Ad Hoc Video Search (Kazuya Ueki)
10:15 – 10:45	Coffee break
10:45 – 11:30	Activity Recognition (Cees Snoek)
11:30 – 12:15	Instance Search (Shin'ichi Satoh)
12:15 – 12:30	General questions and discussion
12:30	Tutorial ends

Tutorial Number: T3

Functional Maps: A Flexible Representation for Learning and Computing Correspondence

Date: Saturday 8th morning

Room: Theresianum 1601

Organizers: Or Litany, Emanuele Rodolà, Maks Ovsjanikov, Leo Guibas

SCHEDULE

08:00 – 08:45	Introduction and Overview
08:50 – 09:35	Computing Functional Maps
09:40 – 10:25	Maps in Shape Collections
10:30 – 11:00	Coffee break
11:00 – 11:45	Learning Correspondence
11:50 – 12:30	Extensions and Applications
12:30	Tutorial ends

Tutorial Number: T6

Normalization Methods for Training Deep Neural Networks: Theory and Practice

Date: Saturday 8th morning

Room: N1189

Organizers: Mete Ozay and Lei Huang

SCHEDULE

08:00 – 08:30	Introduction
08:30 – 09:30	Normalization Techniques Motivation, Methods and Analysis
09:30 – 10:15	Applying Normalization Methods for Computer Vision Tasks in Practise
10:15 – 10:30	Coffee Break
10:30 – 11:15	Mathematical Foundations
11:15 - 12:00	Theoretical Results and Challenges
12:00 – 12:15	Questions and Discussion

Tutorial Number: T7

Representation Learning for Pedestrian Re-identification

Date: Saturday 8th morning

Room: N1080ZG

Organizers: Shengcai Liao, Yang Yang, Liang Zheng

SCHEDULE

09:00 – 09:40	A general introduction and overview of person re-identification
09:40 – 10:40	The seamless corporation of visual descriptors and similarity metrics
10:40 – 11:00	Coffee break
11:00 – 12:20	Deep architectures for representation learning and transfer learning

Tutorial Number: T8

Visual Localization: Feature-based vs. Learned Approaches

Date: Saturday 8th morning

Room: N1070ZG

Organizers: Torsten Sattler, Eric Brachmann

SCHEDULE

09:00	Tutorial starts
09:00 – 09:10	Introduction (Brachmann, Sattler)
09:10 – 10:00	Current State of Feature-based Localization (Sattler)
10:00 – 10:15	Coffee Break
10:15 – 11:15	Current State of Learning-based Localization (Brachmann)
11:15 – 11:30	Coffee Break
11:30 – 12:20	Current Topics & Open Problems (Brachmann, Sattler)
12:20 – 12:40	Questions & Discussion (Brachmann, Sattler)
12:40	Tutorial ends

Tutorial Number: T9

Adversarial Machine Learning

Date: Saturday 8th morning

Room: N1179

Organizers: Battista Biggio, Fabio Roli

SCHEDULE

08:15	Tutorial starts
08:30 – 09:00	Introduction to Adversarial Machine Learning (Fabio Roli)
09:00 – 09:30	Design of Pattern Classifiers in Adversarial Environments (Fabio Roli)
09:30 – 10:30	Machine Learning under attack, Part I (Battista Biggio)
10:30 – 11:00	Coffee break
11:00 – 11:45	Machine Learning under attack, Part II (Battista Biggio)
11:45 – 12:30	Attacks in the Physical World, Summary and Outlook (Fabio Roli)
12:30	Tutorial ends

Tutorial Number: 10

HoloLens as a tool for computer vision research

Date: Saturday 8th morning

Room: Theresianum 606

Organizers: Marc Pollefeys, Shivkumar Swaminathan, Johannes Schoenberger, Andrew Fitzgibbon

SCHEDULE

09:00 – 09:30	Introduction
09:30 – 10:30	HoloLens Research Mode
10:30 – 11:00	Coffee break
11:00 – 11:30	Applications & Demos
11:30 – 12:00	Kinect for Azure Depth Sensor
12:00 – 12:30	Q & A
12:30	Tutorial ends

Tutorial Number: T11

Face Tracking and its Applications

Date: Saturday 8th morning

Room: Theresianum 602

Organizers: Dr. Justus Thies (Technical University of Munich)
Dr. Michael Zollhöfer (Stanford University)
Prof. Dr. Matthias Niessner (Technical University of Munich)
Prof. Dr. Christian Theobalt (Max-Planck-Institute for Informatics)

SCHEDULE

09:00 – 10:20	Basic Topics - Optimization-based Approaches
10:20 – 10:40	Coffee break
10:40 – 12:00	Advanced Topics – Deep-Learning-based Approaches

Workshop Number: 35

Vision Meets Drone: A Challenge

Date: Saturday 8th morning

Room: N1095ZG

Organizers: Pengfei Zhu, Longyin Wen, Xiao Bian, Haibin Ling

SCHEDULE

08:30 – 09:00	Welcome and opening remarks
09:00 – 09:40	Invited Keynote: Direct Visual SLAM for Autonomous Vehicles Keynote Speaker: Daniel Cremers
09:40 – 10:10	Coffee break
10:10 – 10:50	Invited Keynote: Airborne Video Surveillance and Camera Networks Keynote Speaker: Mubarak Shah
10:50 – 11:05	Winner Talk: Hybrid Attention Based Low-Resolution Retina-Net
11:05 – 11:20	Winner Talk : SSD with Comprehensive Feature Enhancement
11:20 – 11:35	Winner Talk : An improved ECO algorithm for preventing camera shake, long-term occlusion and adaptation to target deformation
11:35 – 11:50	Winner Talk: Multi-object tracking with combined constraints and geometry verification
11:50 – 12:30	Awarding ceremony

Workshop Number: 6

2nd International Workshop on Computer Vision for UAVs (UAVision)

Date: Saturday 8th afternoon

Room: N1095ZG

Organizers: Kristof Van Beeck, Toon Goedemé, Tinne Tuytelaars, Davide Scaramuzza

SCHEDULE

13:30	Welcome session
13:40 – 14.00	Teaching UAVs to race Matthias Müller, Vincent Casser, Neil Smith, Dominik Michels, Bernard Ghanem
14:00 – 14:20	Onboard Hyperspectral Image Compression using Compressed Sensing and Deep Learning Saurabh Kumar, Subhasis Chaudhuri, Biplab Banerjee, Feroz Ali
14:20 – 14:40	SafeUAV: Learning to estimate depth and safe landing areas for UAVs from synthetic data Alina E Marcu, Dragos Costea, Vlad Licaret, Mihai Pirvu, Emil Slusanschi, Marius Leordeanu
14:40 – 15:00	Aerial GANeration: Towards Realistic Data Augmentation Using Conditional GAN Stefan Milz
15:00 – 15:45	Coffee break
15:45 – 16:00	Metrics for Real-Time Mono-VSLAM Evaluation including IMU induced Drift with Application to UAV Flight Alexander Hardt-Stremayr, Matthias Schörghuber, Stephan Weiss, Martin Humenberger
16:00 – 16:15	ShuffleDet: Real-Time Vehicle Detection Network in On-board Embedded UAV Imagery Seyed Majid Azimi

16:15 – 16:30	Joint Exploitation of Features and Optical Flow for Real-Time Moving Object Detection on Drones Mehmet Kerim Yücel, Hazal Lezki, Ahu Ozturk, Mehmet Akif Akpinar, Berker Logoglu, Erkut Erdem, Aykut Erdem
16:30 – 16:45	UAV-GESTURE: A Dataset for UAV Control and Gesture Recognition Asanka G Perera, Yee Wei Law, Javaan Chahl
16:45 – 17:00	ChangeNet: A Deep Learning Architecture for Visual Change Detection Ashley Varghese, Jayavardhana Gubbi, Akshaya Ramaswamy, Balamuralidhar P
17:00 – 17:10	Closing session and best paper award
17:10	End of the workshop

Workshop Number: 10

Open Images Challenge Workshop

Date: Saturday 8th afternoon

Room: Theresianum 602

Organizers: Vittorio Ferrari, Alina Kuznetsova, Jordi Pont-Tuset, Matteo Malloci, Jasper Uijlings, Jake Walker, Rodrigo Benenson (Google Research)

SCHEDULE

13:30 - 13:50	Overview of Open Images and the challenge
13:50 - 14:10	Object Detection track - settings, metrics, winners, analysis
14:10 - 15:00	Presentations by three selected object detection participants
15:00 - 15:40	Keynote: A-STAR: Towards Agents that See, Talk, Act, and Reason by Devi Parikh
15:40 - 16:20	Break (posters)
16:20 - 16:40	Visual Relationship Detection track - settings metrics, winners, analysis
16:40 - 17:10	Presentations by two selected VRD participants
17:10 - 17:20	Concluding remarks and plans for future Open Images workshops

Tutorial Number: 13

Instance-level Visual Recognition

Date: Saturday 8th afternoon

Room: N1179

Organizers: Georgia Gkioxari, Ross Girshick, Piotr Dollàr and Kaiming He

SCHEDULE

14:00 – 14:45	Learning Deep Representations for Visual Recognition by Kaiming He
14:45 – 15:30	The Generalized R-CNN Framework for Object Detection by Ross Girshick
15:30 – 16:15	Panoptic Segmentation: Unifying Semantic and Instance Segmentations by Alexander Kirillov
16:15 – 16:45	Coffee Break
16:45 – 17:30	Deep Insights into Convolutional Networks for Video Recognition by Christoph Feichtenhofer
17:30 – 18:15	DensePose: Learning Dense Correspondences in the Wild by Natalia Neverova

Workshop Number: 22

WIDER Face and Pedestrian Challenge

Date: September 8th afternoon

Room: N1090ZG

Organizers: Chen Change Loy, Dahua Lin, Wanli Ouyang, Yuanjun Xiong, Shuo Yang, Qingqiu Huang, Dongzhan Zhou, Wei Xia, Ping Luo, Quanquan Li, Junjie Yan

SCHEDULE

13:30 – 13:35	Opening Remarks
13:35 – 14:10	Invited Talk: Rama Chellappa
14:10 – 14:45	Invited Talk: Sanja Fidler
14:45 – 15:10	Spotlights by WIDER Face Winners
15:10 – 15:35	Spotlights by WIDER Pedestrian Winners
15:35 – 16:00	Coffee Break and Poster Session
16:00 – 16:35	Invited Talk: Shaogang Gong
16:35 – 17:10	Invited Talk: Bernt Schiele
17:10 – 17:35	Spotlights by WIDER Person Search Winners
17:35 – 18:00	Award Ceremony and Closing Remarks

Workshop Number: 41

Vision for XR

Date: Saturday 8th afternoon

Room: Theresianum 601 EG

Organizers: Michael Goesele, Richard Newcombe, Chris Sweeney, Jakob Engel, Julian Straub

SCHEDULE

14:00 – 14:15	Opening remarks -- what's special about XR: VR, AR and MR?
14:15 – 14:45	Physical Reconstruction
14:45 – 15:15	Semantic Reconstruction
15:15 – 15:45	Break
15:45 – 16:15	Self Tracking and SLAM
16:15 – 16:45	Vision on XR Devices
16:45 – 17:15	Keynote
17:15 – 17:45	Panel Discussion
17:45 – 18:00	Closing Remarks
18:00	End of Workshop

Workshop Number: 51

AutoNUE: Autonomous Navigation in Unconstrained Environments

Date: Saturday 8th afternoon

Room: Theresianum 606

Organizers: Manmohan Chandraker (UCSD)
C. V. Jawahar (IIIT Hyderabad)
Anoop Namboodiri (IIIT Hyderabad)
Srikumar Ramalingam (Univ. of Utah)
Anbumani Subramanian (Intel)

SCHEDULE

13:30 – 13:45	Welcome and Background
13:45 – 14:30	Keynote 1: Jitendra Malik (UC Berkeley)
14:30 – 15:15	Dataset and Challenge
15:15 – 16:00	Spotlight and Posters
16:00 – 16:30	Coffee Break
16:30 – 17:15	Keynote 2: Vladlen Koltun (Intel)
17:15 – 18:00	Keynote 3: Andreas Geiger (University of Tübingen)

Workshop Number: 55

Anticipating Human Behavior

Date: Saturday 8th afternoon

Room: N1080ZG

Organizers: Juergen Gall (University of Bonn)
Jan van Gemert (Delft University of Technology)
Kris Kitani (Carnegie Mellon University)

SCHEDULE

13:15 – 13:30 Introduction

13:30 – 14:00 Invited Talk: Abhinav Gupta, Carnegie Mellon University

14:00 – 15:05 **Oral Session**

• Action Anticipation by Predicting Future Dynamic Images
Cristian Rodriguez (Australian National University), Basura Fernando (Australian National University), Hongdong Li (Australian National University)

• Joint Future Semantic and Instance Segmentation Prediction
Camille Couprie (Facebook AI Research), Pauline Luc (Facebook AI Research), Jakob Verbeek (INRIA)

• Context Graph based Video Frame Prediction using Locally Guided Objective
Prateep Bhattacharjee (Indian Institute of Technology Madras), Sukhendu Das (Indian Institute of Technology Madras)

• Predicting Action Tubes
Gurkirt Singh (Oxford Brookes University), Suman Saha (Oxford Brookes University), Fabio Cuzzolin (Oxford Brookes University)

• Forecasting Hands and Objects in Future Frames
Chenyou Fan (Indiana University), Jangwon Lee (Indiana University), Michael S Ryoo (Indiana University)

15:05 – 15:40 Poster Session including Coffee

• When will you do what? - Anticipating Temporal Occurrences of Activities
Yazan Abu Farha (University of Bonn), Alexander Richard (University of Bonn), Juergen Gall (University of Bonn)

· Motion Prediction with Gaussian Process Dynamical Models and Trajectory Optimization
Philipp Kratzer (University of Stuttgart), Marc Toussaint (University of Stuttgart), Jim Mainprice (University of Stuttgart, MPI-IS)

· R2P2: A ReparameteRized Pushforward Policy for Diverse, Precise Generative Path Forecasting
Nicholas Rhinehart (Carnegie Mellon University), Kris M. Kitani (Carnegie Mellon University), Paul Vernaza (Carnegie Mellon University)

· Action Anticipation with RBF Kernelized Feature Mapping RNN
Yuge Shi (Australian National University), Basura Fernando (Australian National University), Richard Hartley (Australian National University)

· Predicting Future Instance Segmentations by Forecasting Convolutional Features
Camille Couprie (Facebook AI Research), Pauline Luc (Facebook AI Research), Yann LeCun (Facebook AI Research), Jakob Verbeek (INRIA)

15:40 – 16:20 Poster Session including Coffee

· Learning to Forecast and Refine Residual Motion for Image-to-Video Generation
Long Zhao (Rutgers University), Xi Peng (Binghamton University), Yu Tian (Rutgers University), Mubbasir Kapadia (Rutgers University), Dimitris Metaxas (Rutgers University)

· Deep Video Generation, Prediction and Completion of Human Action Sequences
Haoye Cai (Hong Kong University of Science and Technology, Stanford University), Chunyan Bai (Hong Kong University of Science and Technology, Carnegie Mellon University), Yu-Wing Tai (Tencent Youtu), Chi-Keung Tang (Hong Kong University of Science and Technology)

· Adversarial Geometry-Aware Human Motion Prediction
Liang-Yan Gui (Carnegie Mellon University), Yu-Xiong Wang (Carnegie Mellon University), Xiaodan Liang (Carnegie Mellon University), José M. F. Moura (Carnegie Mellon University)

· Embarrassingly Simple Model for Early Action Proposal
Marcos Baptista-Ríos (University of Alcalá), Roberto Lopez-Sastre (University of Alcalá), Francisco J. Acevedo-Rodríguez (University of Alcalá), Saturnino Maldonado-Bascon (University of Alcalá)

· Am I done? Predicting Action Progress in Video
Federico Becattini (University of Florence), Lorenzo Seidenari (University of Florence), Tiberio Uricchio (University of Florence), Alberto Del Bimbo (University of Florence), Lamberto Ballan (University of Padova)

· A Novel Semantic Framework for Anticipation of Manipulation Actions
Fatemeh Ziaeetabar (University of Göttingen), Minija Tamosiunaite (University of Göttingen), Florentin Wörgötter (University of Göttingen)

16:20 – 16:50 Invited Talk: Vulnerable Road User Path Prediction Dariu M. Gavrila, Delft University of Technology

16:50 – 17:45 **Oral Session**

· RED: A simple but effective Baseline Predictor for the TrajNet Benchmark
Stefan Becker (Fraunhofer IOSB), Ronny Hug (Fraunhofer IOSB), Wolfgang Hübner (Fraunhofer IOSB), Michael Arens (Fraunhofer IOSB)

· Convolutional Neural Network for Trajectory Prediction
Nishant Nikhil (Indian Institute of Technology Kharagpur), Brendan Morris (University of Nevada Las Vegas)

· Leader's Gaze Behaviour and Alignment of the Action Planing from the Follower's Gaze Cues in Human-Human and Human-Robot Interaction
Nuno Duarte (Instituto Superior Técnico), Mirko Rakovic (Instituto Superior Técnico), Jorge Marques (Instituto Superior Técnico), José Santos-Victor (Instituto Superior Técnico)

· Group LSTM: Group Trajectory Prediction in Crowded Scenarios
Niccolò Bisagno (Università di Trento), Bo Zhang (Dalian Maritime University), Nicola Conci (UNITN)

17:45 – 18:00 Closing Remarks & Discussion

Workshop Number: 64

Brain-Driven Computer Vision

Date: Saturday 8th afternoon

Room: N1189

Organizers: Simone Palazzo, Isaak Kavasidis, Dimitris Kastaniotis, Stavros Dimitriadis

SCHEDULE

13:30 – 13:45	Workshop starts
13:45 – 14:30	Keynote (John Tsotsos, York University)
14:30 – 16:00	Poster session
16:00 – 16:30	Coffee break
16:30 – 17:15	Keynote (Concetto Spampinato, University of Catania)
17:15 – 17:45	Discussion
17:45 – 18:00	Conclusions and future directions
18:00	Workshop ends

Workshop Number: 69

PoseTrack Challenge: Articulated People Tracking in the Wild

Date: September 8th, afternoon

Room: N1070ZG

Organizers: Mykhaylo Andriluka, Umar Iqbal, Christoph Lassner, Eldar Insafutdinov, Leonid Pishchulin, Siyu Tang, Anton Milan, Jürgen Gall, Bernt Schiele

SCHEDULE

13:20 – 13:30	Introduction
13:30 – 14:00	Invited Talk: Christian Theobalt
14:00 – 14:30	Invited Talk: George Papandreou
14:30 – 15:00	Challenge Results
15:00 – 15:45	Poster Session and Coffee Break
15:45 – 16:15	Invited Talk: Cristian Sminchisescu
16:15 – 16:45	Invited Talk: Iasonas Kokkinos
16:45 – 17:15	Closing Remarks & Discussion

Workshop Number: 72

Workshop on Objectionable Content and Misinformation

Date: Saturday 8th, afternoon

Room: Theresianum 1601

Organizers: Cristian Canton, Matthias Niesser, Marius Vlad, Paul Natsev, Mark D. Gianturco, Ruben van der Dussen

SCHEDULE

13:30 – 13:45	Intro from organizers
13:45 – 14:30	Keynote: Michael Zollhöefer (Stanford)
14:30 – 15:00	Orals
15:00 – 14:15	Break
15:15 – 15:45	Poster Session
15:45 – 16:30	Keynote: Luisa Verdoliva (Università degli Studi di Napoli, Federico II)
16:30 – 17:15	Keynote: Alyosha Efros (Berkeley)
17:15 – 18:00	Discussion panel
18:00 – 18:15	Closing Remarks
	Dinner

Tutorial Number: 12

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

Date: Saturday 8th, Full Day

Room: 1200 Carl von Linde

Organizers: Sean Fanello, Julien Valentin, Jonathan Taylor, Christoph Rhemann, Adarsh Kowdle, Jürgen Sturm, Shahram Izadi

SCHEDULE

9:00 – 9:15	Introduction Shahram Izadi Morning Session: Depth Sensors, 3D Capture & Camera Tracking
9:15 – 9:45	Depth Sensors and Algorithms: What, When, Where Sean Fanello
9:45 – 10:15	Triangulation Methods: Basics, Challenges Christoph Rhemann
10:15 – 10:30	Low Compute and Fully Parallel Computer Vision with HashMatch Julien Valentin
10:30 – 11:00	Coffee Break
11:00 – 11:15	Depth Completion of a Single RGB-D Image Yinda Zhang
11:15-11:35	Depth Estimation in the Age of Deep Learning Sameh Khamis
11:35-11:50	Learning a Multi-View Stereo Machine Christian Haene
11:50 – 12:30	Localization and Mapping - ARCore Konstantine Tsotsos
12:30 – 13:30	Lunch Break

Afternoon Session:
World and Human Reconstruction and Understanding

13:30 – 13:45	3D Scene Reconstruction Thomas Funkhouser
13:45 – 14:00	3D Scene Understanding Martin Bokeloh
14:00 – 14:15	Coffee Break
14:15 – 14:45	Non-Linear Optimization Methods Andrea Tagliasacchi
14:45 – 15:00	Parametric Tracking of Hands Anastasia Tkach
15:00 – 15:15	Parametric Tracking of Faces Sofien Bouaziz
15:15 – 15:30	Coffee Break
15:30 – 16:00	Non-Rigid Reconstruction of Humans Kaiwen Guo
16:00 – 16:15	Real-time Compression and Streaming of 4D Performances Danhang Tang
16:15 – 16:30	Neural rendering for Performance Capture Ricardo Martin Brualla
16:30 – 16:45	Machine Learning for Human Motion Understanding on Embedded Devices Pavel Pidlypenskyi

Workshop Number: 42

Workshop on Shortcomings in Vision and Language

Date: Saturday 8th August, Full day

Room: AUDIMAX 0980

Organizers: Raffaella Bernardi, Raquel Fernandez, Spandana Gella, Kushal Kafle, Stefan Lee, Dhruv Batra, and Moin Nabi

SCHEDULE

09:00 – 09:10 Welcome and Opening Remarks

09:10 – 09:50 Invited Talk: Danna Gurari
University of Texas, Austin

09:50 – 10:30 Invited Talk: Aishwarya Agrawal
Georgia Institute of Technology

10:30 – 11:45 **Poster Session 1** (Extended Abstracts) with coffee

P1: Video Object Segmentation with Language Referrin Expressions
Anna Khoreva, Anna Rohrbach, Bernt Schiele

P2: Semantic Action Discrimination in Movie Descriptio Dataset
Andrea Amelio Ravelli, Lorenzo Gregori, Lorenzo Seidenari

P3: The Impact of Words Corpus Stochasticity on Word Spotting in Handwriting Documents
MS Al-Rawi, Dimosthenis Karatzas

P4: Learning to see from experience: But which experience is more propaedeutic?
Ravi Shekhar, Ece Takmaz, Nikos Kondylidis, Claudio Greco, Aashish Venkatesh, Raffaella Bernardi, Raquel Fernandez

P5: Visual Dialogue Needs Symmetry, Goals, and Dynamics: The Example of the MeetUp Task
David Schlangen, Nikolai Ilinykh, Sina Zarrieß

P6: Building Common Ground in Visual Dialogue: The PhotoBook Task and Dataset
Janosch Haber, Raquel Fernandez, Elia Bruni

P7: Entity-Grounded Image Captioning
Annika Lindh, Robert Ross, John Kelleher

P8: Modular Mechanistic Networks for Computational Modelling of Spatial Descriptions
Simon Dobnik, John Kelleher

P9: Visual Question Answering as a Meta Learning Task
Damien Teney, Anton Van Den Hengel

P10: An Evaluative Look at the Evaluation of VQA
Shailza Jolly, Sandro Pezzelle, Tassilo Klein, Moin Nabi

P11: The Visual QA Devil in the Details: The Impact of Early Fusion and Batch Norm on CLEVR
Mateusz Malinowski, Carl Doersch

P12: Make up Your Mind: Towards Consistent Answer Predictions in VQA Models
Arijit Ray, Giedrius Burachas, Karan Sikka, Anirban Roy, Avi Ziskind, Yi Yao, Ajay Divakaran

P13: Visual speech language models
Helen L Bear

P14: Be Different to Be Better: Toward the Integration of Vision and Language
Sandro Pezzelle, Claudio Greco, Aurelie Herbelot, Tassilo Klein, Moin Nabi, Raffaella Bernardi

P15: Towards Speech to Sign Language Translation
Amanda Cardoso Duarte, Gorkem Camli, Jordi Torres, Xavier Giro-i-Nieto

P16: The overlooked role of self-agency in artificial systems
Matthew D Goldberg, Justin Brody, Timothy Clausner, Donald Perlis

P17: Women also Snowboard: Overcoming Bias in Captioning Models
Kaylee Burns, Lisa Anne Hendricks, Kate Saenko, Trevor Darrell, Anna Rohrbach

P18: Estimating Visual Fidelity in Image Captions
Pranava Madhyastha, Josiah Wang, Lucia Specia

P19: Object Hallucination in Image Captioning
Anna Rohrbach, Lisa Anne Hendricks, Kaylee Burns, Trevor Darrell, Kate Saenko

	P20: From entailment to Generation Somayeh jafaritazehjani, Albert Gatt
11:45 – 12:25	Invited Talk: Lucia Specia University of Sheffield
12:25 – 13:20	Spotlight Talks (Full Papers)
12:25 – 12:30	S1: Shortcomings in the Multi-Modal Question Answering Task Monica Haurilet, Ziad Al-Halah, Rainer Stiefelhagen
12:30 – 12:35	S2: Knowing Where to Look? Analysis on Attention of Visual Question Answering System Wei Li, Zehuan Yuan, Changhu Wang
12:35 – 12:40	S3: Pre-gen metrics: Predicting caption quality metrics without generating captions Marc Tanti, Albert Gatt, Adrian Muscat
12:40 – 12:45	S4: Quantifying the amount of visual information used by neural caption generators Marc Tanti, Albert Gatt, Kenneth Camilleri
12:45 – 12:50	S5: Distinctive-attribute Extraction for Image Captioning Boeun Kim, Young Han Lee, Hyedong Jung, Choongsang Cho
12:50 – 12:55	S6: Towards a Fair Evaluation of Zero-Shot Action Recognition from Word Embeddings Alina Roitberg, Manel Martinez, Monica Haurilet, Rainer Stiefelhagen
12:55 – 13:00	S7: How Do End-to-End Image Description Systems Generate Spatial Relations? Mohammad Mehdi Ghanimifard, Simon Dobnik
13:00 – 13:05	S8: How clever is the FiLM model, and how clever can it be? Alexander Kuhnle, Huiyuan Xie, Ann Copestake
13:05 – 13:10	S9: Image-sensitive language modeling for automatic speech recognition Kata Naszadi, Dietrich Klakow
13:10 – 13:15	S10: Improving Context Modelling in Multimodal Dialogue Generation Shubham Agarwal, Ondrej Dusek, Ioannis Konstas, Verena Rieser

13:15 – 13:20 S11: Adding Object Detection Skills to Visual Dialogue Agents
Gabriele Bani, Tim Baumgärtner, Aashish Venkatesh, Davide Belli, Gautier Dagan, Alexander Geenen, Andrii Skliar, Elia Bruni, Raquel Fernandez

13:20 – 14:30 Lunch break

14:30 – 15:30 **Poster Session 2** (Full Papers)

S1: Shortcomings in the Multi-Modal Question Answering Task
Monica Haurilet, Ziad Al-Halah, Rainer Stiefelhagen

S2: Knowing Where to Look? Analysis on Attention of Visual Question Answering System
Wei Li, Zehuan Yuan, Changhu Wang

S3: Pre-gen metrics: Predicting caption quality metrics without generating captions
Marc Tanti, Albert Gatt, Adrian Muscat

S4: Quantifying the amount of visual information used by neural caption generators
Marc Tanti, Albert Gatt, Kenneth Camilleri

S5: Distinctive-attribute Extraction for Image Captioning
Boeun Kim, Young Han Lee, Hyedong Jung, Choongsang Cho

S6: Towards a Fair Evaluation of Zero-Shot Action Recognition from Word Embeddings
Alina Roitberg, Manel Martinez, Monica Haurilet, Rainer Stiefelhagen

S7: How Do End-to-End Image Description Systems Generate Spatial Relations?
Mohammad Mehdi Ghanimifard, Simon Dobnik

S8: How clever is the FiLM model, and how clever can it be?
Alexander Kuhnle, Huiyuan Xie, Ann Copestake

S9: Image-sensitive language modeling for automatic speech recognition
Kata Naszadi, Dietrich Klakow

S10: Improving Context Modelling in Multimodal Dialogue Generation
Shubham Agarwal, Ondrej Dusek, Ioannis Konstas, Verena Rieser

	S11: Adding Object Detection Skills to Visual Dialogue Agents Gabriele Bani, Tim Baumgärtner, Aashish Venkatesh, Davide Belli, Gautier Dagan, Alexander Geenen, Andrii Skliar, Elia Bruni, Raquel Fernandez
15:30 – 16:10	Invited Talk: Vicente Ordóñez Román University of Virginia
16:10 – 17:10	Visual Dialog Challenge (See details @ https://visualdialog.org/)
17:10 – 17:50	Poster Session 3 (Visual Dialog Challenge) with coffee
17:50 – 18:00	Closing Remarks

Workshop Number: 19

Third International Workshop on Egocentric Perception, Interaction and Computing (EPIC)

Date: Sunday 9th, morning

Room: N1090ZG

Organizers: Dima Damen, University of Bristol, UK
Giuseppe Serra, University of Udine, Italy
David Crandall, Indiana University, USA
Giovanni Maria Farinella, University of Catania, Italy
Antonino Furnari, University of Catania, Italy

SCHEDULE

08:20 Welcome

08:30 – 09:05 Keynote: Capturing First-Person alongside Third-Person Videos in the Wild
Abinav Gupta (CMU, USA)

09:05 – 09:35 **Full Paper Presentations**

MACNet: Multi-scale Atrous Convolution Networks for Food Places Classification in Egocentric Photo-streams
Md. Mostafa Kamal Sarker (Rovira i Virgili University), Hatem Rashwan (Rovira i Virgili University), Syeda Furruka Banu (Rovira i Virgili University), Petia Radeva (University of Barcelona) and Domenec Puig (Rovira i Virgili University)

PathGAN: Visual Scanpath Prediction with Generative Adversarial Networks
Marc Assens Reina (Universitat Politecnica de Catalunya), Kevin McGuinness (Dublin City University), Xavier Giro-i-Nieto (Universitat Politecnica de Catalunya), Noel O'Connor (Dublin City University)

09:35 – 09:50 EPIC-KITCHENS 2018 Dataset: Challenges Opening
Will Price (University of Bristol, UK) and Antonino Furnari (University of Catania, Italy)

09:50 – 10:25 Keynote: Video Understanding: Learning More with Less Annotation
Ivan Laptev (INRIA Paris, France)

10:25 – 11:30 Coffee Break and Poster Session

Accepted Abstracts:

Eye Movement Velocity and Gaze Data Generator for Evaluation, Robustness Testing and Assess of Eye Tracking Software and Visualization Tools
Wolfgang Fuhl (University of Tuebingen) and
Enkelejda Kasneci (University of Tuebingen)

Soft-PHOC Descriptor for End-to-End Word Spotting in Egocentric Scene Images
Dena Bazazian (Computer Vision Center),
Dimosthenis Karatzas (Computer Vision Centre) and
Andy Bagdanov (University of Florence)

On the Role of Event Boundaries in Egocentric Activity Recognition from Photo Streams
Alejandro Cartas (University of Barcelona),
Estefanía Talavera (University of Barcelona),
Mariella Dimiccoli (Computer Vision Center) and
Petia Radeva (University of Barcelona)

Joint Attention Estimation from Object-wise 3D Gaze Concurrences in Multiple First-Person Videos and Gazes
Hiroyuki Ishihara (NTT)

Predicting Gaze in Egocentric Video by Learning Task-dependent Attention Transition
Yifei Huang (University of Tokyo),
Minjie Cai (University of Tokyo),
Zhenqiang Li (University of Tokyo) and
Yoichi Sato (University of Tokyo)

A Geometric Model of Spatial Misperception in Virtual Environments
Manuela Chessa (University of Genova) and
Fabio Solari (University of Genova)

Visitors Localization from Egocentric Videos
Francesco Ragusa (University of Catania),
Antonino Furnari (University of Catania),
Sebastiano Battiato (University of Catania),
Giovanni Signorello (University of Catania) and
Giovanni Maria Farinella (University of Catania)

Towards an Embodied Semantic Fovea: Semantic 3D scene reconstruction from ego-centric eye-tracker videos
Mickey Li (Imperial College London),
Pavel Orlov (Imperial College London),
Aldo Faisal (Imperial College London),
Noyan Songur (Imperial College London) and
Stefan Leutenegger (Imperial College London)

The Impact of Temporal Regularisation in Vide Saliency Prediction
Panagiotis Linardos (Universitat Politècnica de Catalunya),
Xavier Giro-i-Nieto (Universitat Politecnica de Catalunya),
Eva Mohedano (Insight Center for Data Analytics),
Monica Cherto (Insight Center for Data Analytics) and
Cathal Gurrin (Dublin City University)

Depth in the Visual Attention Modelling from the Egocentric Perspective of View
Miroslav Laco (Slovak University of Technology in Bratislava) and Wanda Benesova (Slovak University of Technology in Bratislava)

Efficient Egocentric Visual Perception Combining Eye-tracking, a Software Retina and Deep Learning
Jan Paul Siebert (University of Glasgow)

Object-centric Attention for Egocentric Activity Recognition
Swathikiran Sudhakaran (Fondazione Bruno Kessler) and Oswald Lanz (Fondazione Bruno Kessler)

Invited ECCV Papers

Scaling Egocentric Vision: The EPIC-KITCHENS Dataset
Dima Damen (University of Bristol),
Hazel Doughty (University of Bristol),
Giovanni Maria Farinella (University of Catania),
Sanja Fidler (University of Toronto, NVIDIA),
Antonino Furnari (University of Catania),
Evangelos Kazakos (University of Bristol),
Davide Moltisanti (University of Bristol),
Jonathan Munro (University of Bristol),
Toby Perrett (University of Bristol),
Will Price (University of Bristol) and
Michael Wray (University of Bristol)

Joint Person Segmentation and Identification in Synchronized First- and Third-person Videos
Mingze Xu (Indiana University), Chenyou Fan (JD.com), Yuchen Wang (Indiana University), Michael Ryoo (Indiana University) and David Crandall (Indiana University)

Efficient 6-DoF Tracking of Handheld Objects from an Egocentric Viewpoint
Rohit Pandey (Google), Pavel Pidlypenskyi (Google), Shuoran Yang (Google), Christine Kaeser-Chen (Google)

11:30 – 12:05 Keynote: Multimodal and Open-ended Learning
Barbara Caputo (Italian Institute of Technology, Italy)

12:05 – 12:35 **Full Paper Presentations II**

Leveraging Uncertainty to Rethink Loss Functions and Evaluation Measures for Egocentric Action Anticipation
Antonino Furnari (University of Catania), Sebastiano Battiato (University of Catania) and Giovanni Maria Farinella (University of Catania)

MAM: Transfer Learning for Fully Automatic Video Annotation and Specialized Detector Creation
Wolfgang Fuhl (University of Tuebingen), Nora J Castner (University Tübingen), Markus Holzer (Bosch), Lin Zhuang (Bosch) and Enkelejda Kasneci (University of Tuebingen)

12:35 – 12:40 Concluding Remarks

12:40 Workshop Ends

Workshop Number: 21

9th International Workshop on Human Behavior Understanding: Towards Generating Realistic Visual Data of Human Behavior

Date: Sunday 9th, morning

Room: AUDIMAX 0980

Organizers: Xavier Alameda-Pineda, Elisa Ricci, Albert Ali Salah, Nicu Sebe, Shuicheng Yan

SCHEDULE

08:30 – 08:45	Welcome and general remarks
08:45 – 09:30	Keytone speaker: Stefanos Zafeiriou (Imperial College London)
09:30 – 09:45	Oral presentation #1
09:45 – 10:00	Oral presentation #2
10:00 – 10:15	Oral presentation #3
10:15 – 10:30	Oral presentation #4
10:30 – 11:30	Poster Session & Coffee Break
11:30 – 12:15	Keynote Speaker: Michael Black (Max Planck Institute & Amazon)
12:15	Conclusion

Workshop Number: 38

4th International Workshop on Recovering 6D Object Pose

Date: Sunday 9th, morning

Room: Theresianum 606

Organizers: Rigas Kouskouridas, Tomas Hodan, Krzysztof Walas, Tae-Kyun Kim, Jiri Matas, Carsten Rother, Frank Michel, Vincent Lepetit, Ales Leonardis, Carsten Steger, Caner Sahin

SCHEDULE

08:00 – 08:10	Workshop starts
08:10 – 08:40	Invited talk 1, Federico Tombari (TU Munich)
08:40 – 09:10	Invited talk 2, Kostas Bekris (Rutgers University)
09:10 – 09:50	Oral presentations of selected workshop papers
09:50 – 10:20	Coffee break
10:20 – 10:50	Invited talk 3, Kurt Konolige (X Robotics)
10:50 – 11:20	Invited talk 4, Thibault Groueix (Ecole Nationale des Ponts et Chaussées)
11:20 – 11:40	BOP: Benchmark for 6D Object Pose Estimation, Tomas Hodan (CTU in Prague)
11:40 – 12:30	Poster session (accepted papers, extended abstracts, invited posters)
12:30	Workshop ends

Workshop Number: 60

Multimodal Learning and Applications Workshop

Date: Sunday 9th, morning

Room: Theresianum 602

Organizers: Paolo Rota, Vittorio Murino
Michael Yang, Bodo Rosenhahn

SCHEDULE

08:20 – 08:30	Initial remarks and workshop introduction
08:30 – 08:50	(ORAL1) - Boosting LiDAR-based Semantic Labeling by Cross-Modal Training Data Generation Florian Piewak; Peter Pinggera; Manuel Schäfer; David Peter; Beate Schwarz; Nick Schneider; Markus Enzweiler; David Pfeiffer; Marius Zöllner
08:50 – 09:40	Invited Speaker: Daniel Cremers
09:40 – 10:00	(ORAL2) - Visually Indicated Sound Generation by Perceptually Optimized Classification Kan Chen; Chuanxi Zhang; Chen Fang; Zhaowen Wang; Trung Bui; Ram Nevatia
10:00 – 10:20	(ORAL3) - Learning to Learn from Web Data through Deep Semantic Embeddings Raul Gomez; Lluis Gomez; Jaume Gibert; Dimosthenis Karatzas
10:20 – 10:30	Coffee Break
10:30 – 11:20	Invited Speaker: Raquel Urtasun
11:20 – 12:00	Spotlight session (3 mins presentation for each poster)
12:00 – 12:50	Poster Session
12:50	Closing

Workshop Number: 67

Second Workshop on 3D Reconstruction Meets Semantics (3DRMS)

Date: Sunday 9th, morning

Room: N1179

Organizers: Radim Tylecek, Torsten Sattler, Thomas Brox, Marc Pollefeys, Robert B. Fisher, Theo Gevers

SCHEDULE

09:00 – 09:05 Introduction by the organizers

09:05 – 09:35 Invited Talk: Andrew Davison

09:35 – 10:15 **Spotlight presentations of accepted papers and extended abstracts**

A Deeper Look at 3D Shape Classifiers
Jong-Chyi Su, Matheus Gadelha, Rui Wang, Subhransu Maji

3D-PSRNet: Part Segmented 3D Point Cloud Reconstruction from a Single Image
Priyanka Mandikal, Navaneet K L, R Venkatesh Babu,

Exploiting Multi-Layer Features Using a CNN-RNN Approach for RGB-D Object Recognition
Ali Caglayan, Ahmet Burak Can

Temporally Consistent Depth Estimation in Videos with Recurrent Architectures
Denis Tananaev, Huizhong Zhou, Benjamin Ummenhofer, Thomas Brox

End-to-end 6-DoF Object Pose Estimation through Differentiable Rasterization
Andrea Palazzi, Luca Bergamini, Simone Calderara, Rita Cucchiara

YOLO3D: End-to-end real-time 3D Oriented Object Bounding Box Detection from LiDAR Point Cloud
Mohamed Zahran, Ahmad ElSallab, Waleed Ali, Sherif Abdelkarim, Mahmoud Zidan

Increasing the robustness of CNN-based human body segmentation in range images by modeling sensor-specific artifacts
Lama Seoud, Jonathan Boisvert, Marc-Antoine Drouin, Michel Picard, Guy Godin

Future Semantic Segmentation with 3D Structure (extended abstract)
Suhani Vora, Anelia Angelova, Soeren Pirk, Reza Majhourian

PanoRoom: From the Sphere to the 3D Layout (extended abstract)
Clara Fernandez Labrador, José María Fácil, Alejandro Perez Yus, Cedric Demonceaux, Josechu Guerrero

10:15 – 10:45	Invited Talk: Thomas Funkhouser
10:45 – 11:30	Coffee Break & Poster presentations
11:30 – 11:40	Discussion of the challenge and results
11:40 – 10:50	Oral presentations by challenge winner
11:50 – 12:20	Invited Talk: Christian Häne
12:20 – 12:40	Panel Discussion with Invited Speakers

Workshop Number: 2

6th Workshop on Assistive Computer Vision and Robotics

Date: Sunday 9th, afternoon

Room: N1179

Organizers: Giovanni Maria Farinella, Marco Leo, Gerard Medioni, Mohan Trivedi

SCHEDULE

13:30 – 13:45	Workshop introduction by the general chairs
13:45 – 14:30	Invited Talk by Tae-Kyun Kim (Imperial College of London, United Kingdom) "3D Hand Pose Estimation for Novel Man-Machine Interface"
14:30 – 15:10	Oral Session 1
	22 Deep execution monitor for robot assistive tasks Fiora Pirri (University of Rome, Sapienza)*; Lorenzo Mauro (ALCOR Lab, Sapienza.); Edoardo Alati (ALCOR Lab, Sapienza.); Gianluca Massimiani (ALCOR Lab, Sapienza.); Marta Sanzari (Sapienza University of Rome); Valsamis Ntouskos (Sapienza University of Rome)
	18 Personalized indoor navigation via multimodal sensors and high-level semantic information Vishnu Nair (The City College of New York)*; Manjekar Budhai (The City College of New York); Greg Olmschenk (CUNY Graduate Center); William H. Seiple (Lighthouse Guild); Zhigang Zhu (CUNY City College and Graduate Center)
15:10 – 15:30	Coffee Break
15:30 – 16:15	Invited Talk by Fabio Galasso (OSRAM GmbH, Germany) "Computer Vision and Smart Lighting relevant to Assistive Technologies"
16:15 – 16:55	Oral Session 2

20 Comparing methods for assessment of facial dynamics in patients with major neurocognitive disorders
Yaohui WANG (INRIA)*; Antitza Dantcheva (INRIA); Jean-Claude Broutart (GSF Noisiez); Philippe Robert (EA CoBTeK – University Cote d'Azur); Francois Bremond (Inria Sophia Antipolis, France); Piotr Bilinski (University of Oxford)

24 Chasing feet in the wild: A proposed egocentric motion-aware gait assessment tool
Mina Nouredanesh (University of Waterloo)*; James Tung (University of Waterloo)

16:55 – 17:55 Poster Session (it will include also poster of papers presented in oral session 1 and 2)

11 Computer Vision for Medical Infant Motion Analysis: State of the Art and RGB-D Data Set
Nikolas Hesse (Fraunhofer IOSB)*; Christoph Bodensteiner (Fraunhofer IOSB); Michael Arens (Fraunhofer IOSB); Ulrich Hofmann (University Medical Center Freiburg); Raphael Weinberger (Dr. v. Hauner Children's Hospital, University of Munich (LMU)); Sebastian Schroeder (Dr. v. Hauner Children's Hospital, University of Munich (LMU))

13 Human-computer interaction approaches for the assessment and the practice of the cognitive capabilities of elderly people
Manuela Chessa (University of Genova, Italy)*; Chiara Bassano (Univeristy of Genova); Elisa Gusai (University of Genova); Alice Evelina Martis (University of Genova); Fabio Solari (University of Genova, Italy)

16 An empirical study towards understanding how deep convolutional nets recognize falls
Yan Zhang (Institute of Neural Information Processing, Ulm University)*; Heiko Neumann (Ulm University)

6 Recovering 6D Object Pose: Reviews and Multi-modal Analyses
Caner Sahin (Imperial College London)*; Tae-Kyun Kim (Imperial College London)

25 Inferring Human Knowledgeability from Eye Gaze in M-learning Environments
Oya Celiktutan (Imperial College London)*; Yiannis Demiris (Imperial College London)

12 Vision Augmented Robot Feeding
Alexandre Candeias (Instituto Superior Tecnico)*; Travers Rhodes (Carnegie Mellon University); Manuel Marques (Instituto Superior Tecnico, Portugal); Joao Paulo Costeira (Instituto Superior Tecnico); Manuela Veloso (Carnegie Mellon University)

14 Analysis of the Effect of Sensors for End-to-End Machine Learning Odometry
Carlos Marquez (Intel)*; Dexmont Pena (Intel)

15 RAMCIP Robot: A Personal Robotic Assistant; Demonstration of a Complete Framework
Ioannis Kostavelis (Center for Research and Technology, Hellas, Information & Technologies Institute)*; Dimitrios Giakoumis (Center for Research and Technology, Hellas, Information & Technologies Institute); Georgia Peleka (Centre for Research and Technology, Hellas, Information Technologies Institute); Andreas Kargakos (Centre for Research and Technology, Hellas, Information Technologies Institute); Evangelos Skartados (Centre for Research and Technology, Hellas, Information Technologies Institute); Manolis Vasileiadis (Centre for Research and Technology, Hellas, Information Technologies Institute); Dimitrios Tzovaras (Centre for Research and Technology Hellas)

17:55 – 18:00 Closing and remarks by the general chairs

Workshop Number: 7

4th International Workshop on Observing and Understanding Hands in Action (HANDS2018)

Date: Sunday 9th, afternoon

Room: Theresianum 606

Organizers: Tae-Kyun Kim, Guillermo Garcia-Hernando, Antonis Argyros, Vincent Lepetit, Iason Oikonomidis, Angela Yao

SCHEDULE

13:30 – 13:45	Welcome and introduction
13:45 – 14:15	Christian Theobalt (MPII, Saarland University)
14:15 – 14:45	Tamim Asfour (KIT)
14:45 – 15:25	Andrew Fitzgibbon (Microsoft)
15:25 – 16:15	Poster session and coffee break

Accepted papers:

• Hand-tremor frequency estimation in videos. Silvia L Pintea (TUDelft); Jian Zheng; Xilin Li; Paulina J.M. Bank; Jacobus J. van Hilten; Jan van Gemert.

• DrawInAir: A Lightweight Gestural Interface Based on Fingertip Regression. Gaurav Garg (TCS Research); Srinidhi Hegde; Ramakrishna Perla; Ramya Hebbalaguppe.

• Adapting Egocentric Visual Hand Pose Estimation Towards a Robot-Controlled Exoskeleton. Gerald Baulig (Reutlingen University); Thomas Gulde; Cristobal Curio.

• Estimating 2D Multi-Hand Poses From Single Depth Images. Le Duan (University of Konstanz); Minmin Shen; Song Cui; Zhexiao Guo; Oliver Deussen

• Spatial-Temporal Attention Res-TCN for Skeleton-based Dynamic Hand Gesture Recognition. Jingxuan Hou (Tsinghua University); Guijin Wang; Xinghao Chen; Jing-Hao Xue; Rui Zhu; Huazhong Yang.

• Task Oriented Hand Motion Retargeting for Dexterous Manipulation Imitation. Dafni Antotsiou (Imperial College London); Guillermo Garcia-Hernando; Tae-Kyun Kim.

Extended abstracts:

• Model-based Hand Pose Estimation for Generalized Hand Shape with Spatial Transformer Network. Shile Li (TUM); Jan Wöhlke; Dongheui Lee.

• A new dataset and human benchmark for partially-occluded hand-pose recognition during hand-object interactions from monocular RGB images. Andrei Barbu (MIT); Battushig Myanganbayar; Cristina Mata; Gil Dekel; Guy Ben-Yosef; Boris Katz.

• 3D Hand Pose Estimation from Monocular RGB Images using Advanced Conditional GAN. Le Manh Quan (Sejong University); Nguyen Hoang Linh; Yong-Guk Kim.

Invited posters:

• HandMap: Robust Hand Pose Estimation via Intermediate Dense Guidance Map Supervision. Xiaokun Wu (University of Bath); Daniel Finnegan; Eamonn O'Neill; Yong-Liang Yang.

• Point-to-Point Regression PointNet for 3D Hand Pose Estimation. Liuhao Ge (NTU); Zhou Ren; Junsong Yuan.

• Joint 3D tracking of a deformable object in interaction with a hand. Aggeliki Tsoli (FORTH); Antonis Argyros.

• Occlusion-aware Hand Pose Estimation Using Hierarchical Mixture Density Network. Qi Ye (Imperial College London); Tae-Kyun Kim.

• Hand Pose Estimation via Latent 2.5D Heatmap Regression. Umar Iqbal (University of Bonn); Pavlo Molchanov; Thomas Breuel; Jürgen Gall; Kautz Jan.

• Weakly-supervised 3D Hand Pose Estimation from Monocular RGB Images. Yujun Cai (NTU), Liuhao Ge, Jianfei Cai, Junsong Yuan.

16:15 – 16:45	Robert Wang (Oculus Research)
16:45 – 17:15:	TBD
17:15 – 17:30	Conclusion and prizes (best paper and best poster award).

Workshop Number: 16

1st Person in Context (PIC) Workshop and Challenge

Date: Sunday 9th, afternoon

Room: Theresianum 602

Organizers: Si Liu, Jiashi Feng, Jizhong Han, Shuicheng Yan

SCHEDULE

13:30 – 13:40	Opening remarks, The Person In Context (PIC) challenge introduction and results
13:40 – 13:50	Oral talk1: Winner of PIC challenge
13:50 – 14:00	Oral talk2: Runner-up/third place of PIC challenge
14:00 – 14:30	Invited talk 1: Bernt Schiele, Max-Planck-Institut für Informatik
14:30 – 15:00	Invited talk 2: Ming-Hsuan Yang, Professor, University of California at Merced
15:00 – 15:30	Invited talk 3: Alan Yuille, Professor, Johns Hopkins University
15:30 – 16:00	Invited talk 4: Jia Deng, Assistant Professor, University of Michigan
16:00 – 16:30	Invited talk 5: Wenjun Zeng, Microsoft Research Asia
16:30 – 17:00	Invited talk 5: Tao Mei, Technical Vice President of JD.com
17:00 – 17:30	Invited talk 6: Changhu Wang, Technical Director of outiao AI Lab
17:30 – 17:35	Awards & Future Plans

Workshop Number: 27

360° Perception and Interaction

Date: Sunday 9th, afternoon

Room: AUDIMAX 0980

Organizers: Min Sun, Yu-Chuan Su, Wei-Sheng Lai, Liwei Chan, Hou-Ning Hu, Silvio Savarese, Kristen Grauman, Ming-Hsuan Yang

SCHEDULE

13:20 – 13:30	Opening remark
13:30 – 14:00	Invited talk VR Video, Steve Seitz (University of Washington)
14:00 – 14:30	Invited talk 360° Video for Robot Navigation, Marc Pollefeys (ETH Zurich)
14:30 – 15:00	Coffee break
15:00 – 15:30	Invited talk VR Video Editing Tools, Aaron Hertzmann (Adobe Research)
15:30 – 16:00	Invited talk Measurable 360, Shannon Chen (Facebook)
16:00 – 17:30	**Posters** Saliency Detection in 360° Videos Ziheng Zhang, Yanyu Xu, Jingyi Yu, Shenghua Gao Gaze Prediction in Dynamic 360° Immersive Video Yanyu Xu, Yanbing Dong, Junru Wu, Zhengzhong Sun, Zhiru Shi, Jingyi Yu, Shenghua Gao Self-Supervised Learning of Depth and Camera Motion from 360° Videos Fu-En Wang, Hou-Ning Hu, Hsien-Tzu Cheng, Juan-Ting Lin, Shang-Ta Yang, Meng-Li Shih, James Hung-Kuo Chu, Min Sun

A Memory Network Approach for Story-based Temporal Summarization of 360° Videos
Sangho Lee, Jinyoung Sung, Youngjae Yu, Gunhee Kim

Deep Learning-based Human Detection on Fisheye Images
Hsueh-Ming Hang, Shao-Yi Wang

Eliminating the Blind Spot: Adapting 3D Object Detection and Monocular Depth Estimation to 360° Panoramic Imagery
Gregoire Payen de La Garanderie,
Amir Atapour-Abarghouei, Toby Breckon

Towards 360° Show-and-Tell
Shih-Han Chou, Yi-Chun Chen, Cheng Sun,
Kuo-Hao Zeng, Ching Ju Cheng, Jianlong Fu, Min Sun

360D: A dataset and baseline for dense depth estimation from 360 images
Antonis Karakottas, Nikolaos Zioulis, Dimitrios Zarpalas,
Petros Daras

Binocular Spherical Stereo Camera Disparity Map Estimation and 3D View-synthesis
Hsueh-Ming Hang, Tung-Ting Chiang, Wen-Hsiao Peng

PathGAN: Visual Scanpath Prediction with Generative Adversarial Networks
Marc Assens Reina, Kevin McGuinness, Xavier Giro-i-Nieto,
Noel O'Connor

Labeling Panoramas With Spherical Hourglass Networks,
Carlos Esteves, Kostas Daniilidis, Ameesh Makadia

The Effect of Motion Parallax and Binocular Stereopsis on Visual Comfort and Size Perception in Virtual Reality,
Jayant Thatte, Bernd Girod

17:30 – 17:40 Closing remark

Workshop Number: 66

5th Women in Computer Vision Workshop

Date: Sunday 9th, afternoon

Room: N1090ZG

Organizers: Zeynep Akata, Dena Bazazian, Yana Hasson, Angjoo Kanazawa, Hildegard Kuehne, Gul Varol

SCHEDULE

13:00 – 13:20	Lunch bags in the poster area
13:20 – 13:30	Introduction
13:30 – 14:00	Keynote: Vision & Language by Tamara Berg (UNC Chapel Hill, Shopagon)
14:00 – 14:15	**Oral session 1:**
	Deep Video Color Propagation by Simone Meyer (ETH Zurich)
	Fashion is Taking Shape: Understanding Clothing Preference based on Body Shape from Online Sources by Hosnieh Sattar (MPI and Saarland University)
	Unsupervised Learning and Segmentation of Complex Activities from Video by Fadime Sener (University of Bonn)
14:15 – 14:45	Keynote: Adapting Neural Networks to New Tasks by Svetlana Lazebnik (University of Illinois at Urbana-Champaign)
14:45 – 16:20	Poster session and coffee break
16:20 – 16:50	Keynote: Explainable AI Models and Why We Need Them by Kate Saenko (Boston University)
16:50 – 17:00	**Oral session 2:**
	Tracking Extreme Climate Events by Sookyung Kim (Lawrence Livermore National Laboratory)
	A Deep Look into Hand Segmentation by Aisha Urooj (University of Central Florida)

17:00 – 17:50	Panel session: Tamara Berg (UNC Chapel Hill, Shopagon), Andrew Fitzgibbon (Microsoft HoloLens), Svetlana Lazebnik (University of Illinois at Urbana-Champaign), Kate Saenko (Boston University), Bernt Schiele (MPI)
17:50 – 18:00	Closing remarks and prizes

Workshop Number: 37

Joint COCO and Mapillary Recognition Challenge Workshop

Date: Sunday 9th, full day

Room: N1080ZG

Organizers:

COCO steering committee:
Tsung-Yi Lin (Google Brain)
Genevieve Patterson (Microsoft Research)
Matteo R. Ronchi (Caltech)
Yin Cui (Cornell)
Iasonas Kokkinos (Facebook AI research)
Michael Maire (TTI-Chicago)
Serge Belongie (Cornell)
Lubomir Bourdev (WaveOne, Inc.)
Ross Girshick (Facebook AI Research)
James Hays (Georgia Tech)
Pietro Perona (Caltech)
Deva Ramanan (CMU)
Larry Zitnick (Facebook AI Research)
Piotr Dollár (Facebook AI Research)

Mapillary steering committee:
Samuel Rota Bulò (Mapillary Research)
Lorenzo Porzi (Mapillary Research)
Peter Kontschieder (Mapillary Research)

Additional organizers:
Alexander Kirillov (Heidelberg University)
Holger Caesar (University of Edinburgh)
Jasper Uijlings (Google Research)
Vittorio Ferrari (University of Edinburgh and Google Research)

SCHEDULE

9:00 – 9:30	Opening remarks
9:30 – 10:30	COCO Instance Segmentation Challenge
10:30 – 11:00	Coffee + Posters
11:00 – 12:00	COCO Panoptic Segmentation Challenge
12:00 – 13:30	Lunch
13:30 – 14:00	COCO Person Keypoints Challenge

14:00 – 14:30	COCO DensePose Challenge
15:00 – 15:30	Coffee + Posters
14:30 – 15:00	Mapillary Instance Segmentation Challenge
15:30 – 16:00	Mapillary Panoptic Segmentation Challenge
16:00 – 16:30	Andreas Geiger: TBD
16:30 – 17:00	Future Plans & Discussion
17:00 – 18:00	Posters

Workshop Number: 39-45

Visual Learning and Embodied Agents in Simulation Environments

Date: Sunday 9th, full day

Room: N1095ZG

Organizers: Peter Anderson, Manolis Savva, Angel X. Chang, Saurabh Gupta, Amir R. Zamir, Stefan Lee, Samyak Datta, Li Yi, Hao Su, Qixing Huang, Cewu Lu, Leonidas Guibas

SCHEDULE

08:45 – 09:00	Welcome and Introduction
09:00 – 09:25	Invited Talk: Peter Welinder
09:25 – 09:50	Invited Talk: Alan Yuille
09:50 – 10:15	Invited Talk: Sanja Fidler
10:15 – 11:00	Coffee and Posters / Demos
11:00 – 11:25	Invited Talk: Boqing Gong
11:25 – 11:50	Invited Talk: Lawson Wong
11:50 – 12:10	Poster Spotlight Presentations
12:10 – 12:30	Simulation Environment Spotlights
12:30 – 13:30	Lunch break
13:30 – 13:55	Invited Talk: Abhinav Gupta
13:55 – 14:20	Invited Talk: Anton van den Hengel
14:20 – 14:45	Invited Talk: Vladlen Koltun
14:45 – 15:30	Coffee and Posters / Demos
15:30 – 15:55	Invited Talk: Raia Hadsell
15:55 – 16:20	Invited Talk: Dhruv Batra
16:20 – 16:45	Invited Talk: Jitendra Malik
16:45 – 17:30	Panel Discussion

Workshop Number: 46

2nd Workshop on Youtube-8M Large-Scale Video Understanding

Date: Sunday 9th, full day

Room: N1070ZG

Organizers: Apostol (Paul) Natsev, Rahul Sukthankar, Joonseok Lee, George Toderici

SCHEDULE

9:00 – 9:05	Opening Remarks
9:05 – 9:30	Overview of YouTube-8M Dataset, Challenge
9:30 – 10:00	Invited Talk 1: Andrew Zisserman
10:00 – 10:30	Invited Talk 2: Rene Vidal
10:30 – 10:45	Coffee Break
10:45 – 12:00	Oral Session 1
12:00 – 13:00	Lunch
13:00 – 13:30	Invited Talk 3: Josef Sivic
13:30 – 14:00	Invited Talk 4: Manohar Paluri
14:00 – 14:30	YouTube-8M Classification Challenge Summary, Organizers' Lightning Talks
14:30 – 15:45	Poster Session + Coffee Break
15:45 – 17:00	Oral Session 2
17:00 – 17:20	Closing and Award Ceremony

Workshop Number: 50

11th POCV Workshop: Action, Perception and Organization

Date: Sunday 9th, full day

Room: Theresianum 601 RG

Organizers: Deepak Pathak, Bharath Hariharan

SCHEDULE

09:00	Workshop starts
09:15 – 09:50	Talk TBD
09:55 – 10:30	Talk TBD
10:30 – 10:45	Coffee break
10:45 – 11:15	Spotlight
11:15 – 11:50	Talk TBD
11:55 – 13:30	Lunch
13:30 – 14:05	Talk TBD
14:05 – 14:40	Talk TBD
14:40 – 15:15	Talk TBD
15:15 – 15:45	Coffee break
15:45 – 16:20	Talk TBD
16:20 – 16:55	Talk TBD
17:00 – 18:00	Posters
18:00	Workshop ends

Workshop Number: 52

ApolloScape: Vision-based Navigation for Autonomous Driving

Date: Sunday 9th, full day

Room: 1200 Carl von Linde Hall

Organizers: Peng Wang, Baidu Research Institution
Ruigang Yang, Baidu Research /University of Kentucky
Andreas Geiger, UMPI/ETH-Zurich
Hongdong Li, Australian National University
Alan Yuille, John Hopkins University

SCHEDULE

08:30 – 08:40	Workshop starts, Workshop Organizers
08:40 – 09:20	Invited talk, Andreas Geiger (MPI)
09:20 – 10:00	Invited talk, Raquel Urtasun (Uber/ University of Toronto)
10:00 – 10:20	Coffee break
10:20 – 11:00	Invited talk, Liang Wang (Baidu)
11:10 – 11:50	Invited talk, Aurora (tentative)
11:50 – 13:30	Lunch break
13:30 – 14:10	Invited talk, Hao Su (UC-San Diego)
14:10 – 14:30	Challenge Award Ceremony
14:30 – 15:30	Award Presentations
15:30 – 16:30	Coffee break + Poster Sessions
16:30 – 17:10	Invited talk, NVIDIA (tentative)
17:10 – 17:50	Invited talk, Marc Pollefeys (ETH-Zurich)
17:50 – 18:00	Concluding Remarks

Workshop Number: 53

VISART

Date: Sunday 9th, full day

Room: Theresianum 1601

Organizers: Stuart James, Leonardo Impett, Peter Hall, Joao Paulo Costeira, Peter Bell, Alessio Del Bue

SCHEDULE

9:00 – 9:15 Welcome remarks

9:15 – 10:00 Invited Talk: "The Art of Vision" Björn Ommer

10:00 – 11:00 **S1: Deep in Art**

"What was Monet seeing while painting? Translating artworks to photo-realistic images"
Matteo Tomei, Lorenzo Baraldi, Marcella Cornia, Rita Cucchiara

"Weakly Supervised Object Detection in Artworks"
Nicolas Gonthier, Yann Gousseau, Saïd Ladjal, Olivier Bonfait

"Deep Transfer Learning for Art Classification Problems"
Matthia Sabatelli, Mike Kestemont, Walter Daelmans, Pierre Geurts

11:00 – 11:30 Coffee Break

11:30 – 12:15 Invited Talk: "Deep Interdisciplinary Learning: Computer Vision and Art History"
Peter Bell

12:15 – 13:00 **S2: Reflections and Tools**

"A Reflection on How Artworks Are Processed and Analyzed by Computer Vision Author"
Sabine Lang, Björn Ommer

"A Digital Tool to Understand the Pictorial Procedures of 17th century Realism"
Francesca Di Cicco, Lisa Wiersma, Maarten Wijntjes, Joris Dik, Jeroen Stumpel, Sylvia Pont

	"Images of Image Machines. Visual Interpretability in Computer Vision for Art" Fabian Offert
13:00 – 14:00	Lunch
14:30 – 15:15	Invited Talk: "Sketching with Style: Visual Search with Sketches and Aesthetic Context" John Collomosse
15:15 – 15:30	Coffee Break
15:30 – 16:50	**S3: Interpreting and Understanding**
	"Seeing the World Through Machinic Eyes: Reflections on Computer Vision in the Arts" Marijke Goeting
	"Saliency-driven Variational Retargeting for Historical Map" Filippo Bergamasco, Arianna Traviglia, Andrea Torsello
	"How to Read Paintings: Semantic Art Understanding with Multi-Modal Retrieval" Noa Garcia, George Vogiatzis
	"Analyzing Eye Movements using Multi-Fixation Pattern Analysis with Deep Learning" Sanjana Kapisthalam, Christopher Kanan, Elena Fedorovskaya
16:50 – 17:05	Coffee Break
17:05 – 17:50	Invited Talk: "On the Limits and Potentialities of Deep Learning for Cultural Analysis" Matteo Pasquinelli
17:50 – 18:00	Closing Remarks
18:00 – 21:00	Social Event

Workshop Number: 61

2nd International Workshop on Compact and Efficient Feature Representation and Learning in Computer Vision (CEFRL 2018)

Date: Sunday 9th, full day

Room: N1189

Organizers: Jie Qin, Li Liu, Li Liu, Fan Zhu, Matti Pietikäinen, Luc Van Gool

SCHEDULE

09:00 –	Workshop starts
09:00 – 09:05	Welcome introduction
09:05 – 09:40	Invited talk 1
09:40 – 10:15	Invited talk 2
10:15 – 10:40	Coffee break
10:40 – 12:00	Oral session 1
12:00 – 14:00	Lunch break
14:00 – 14:20	Invited talk 3
14:20 – 15:20	Poster session
15:20 – 16:40	Oral session 2
16:40 – 16:50	Award ceremony
16:50 – 17:00	Closing remarks
17:00 –	Workshop ends

1A - Learning for Vision 1	September 10, 08:30 am – 09:45 am **Chairs:** Andrea Vedaldi, Oxford Timothy Hospedales, University of Edinburgh	
	Title	Authors
O-1A-01	Convolutional Networks with Adaptive Computation Graphs	Andreas Veit*, Cornell University; Serge Belongie, Cornell University
O-1A-02	Progressive Neural Architecture Search	Chenxi Liu*, Johns Hopkins University; Maxim Neumann, Google; Barret Zoph, Google; Jon Shlens, Google; Wei Hua, Google; Li-Jia Li, Google; Li Fei-Fei, Stanford University; Alan Yuille, Johns Hopkins University; Jonathan Huang, Google; Kevin Murphy, Google
O-1A-03	Diverse Image-to-Image Translation via Disentangled Representations	Hsin-Ying Lee*, University of California, Merced; Hung-Yu Tseng, University of California, Merced; Maneesh Singh, Verisk Analytics; Jia-Bin Huang, Virginia Tech; Ming-Hsuan Yang, University of California at Merced
O-1A-04	Lifting Layers: Analysis and Applications	Michael Moeller*, University of Siegen; Peter Ochs, Saarland University; Tim Meinhardt, Technical University of Munich; Laura Leal-Taixé, TUM
O-1A-05	Learning with Biased Complementary Labels	Xiyu Yu*, The University of Sydney; Tongliang Liu, The University of Sydney; Mingming Gong, University of Pittsburgh; Dacheng Tao, University of Sydney

Poster Session 1A	September 10, 10:00 am – 12:00 pm, see page 142

Demo Session 1A	September 10, 10:00 am – 12:00 pm, see page 96

Lunch	12:00 pm – 01:00 pm

1B - Computational Photography 1	September 10, 01:00 pm – 02:15 pm Chairs: Jan-Michael Frahm, University of North Carolina at Chapel Hill Gabriel Brostow, University College London	
	Title	Authors
O-1B-01	Light Structure from Pin Motion: Simple and Accurate Point Light Calibration for Physics-based Modeling	Hiroaki Santo*, Osaka University; Michael Waechter, Osaka University; Masaki Samejima, Osaka University; Yusuke Sugano, Osaka University; Yasuyuki Matsushita, Osaka University
O-1B-02	Programmable Light Curtains	Jian Wang*, Carnegie Mellon University; Joe Bartels, Carnegie Mellon University; William Whittaker, Carnegie Mellon University; Aswin Sankaranarayanan, Carnegie Mellon University; Srinivasa Narasimhan, Carnegie Mellon University
O-1B-03	Learning to Separate Object Sounds by Watching Unlabeled Video	Ruohan Gao*, University of Texas at Austin; Rogerio Feris, IBM Research; Kristen Grauman, University of Texas
O-1B-04	Coded Two-Bucket Cameras for Computer Vision	Mian Wei, University of Toronto; Navid Navid Sarhangnejad, University of Toronto; Zhengfan Xia, University of Toronto; Nikola Katic, University of Toronto; Roman Genov, University of Toronto; Kyros Kutulakos*, University of Toronto
O-1B-05	Materials for Masses: SVBRDF Acquisition with a Single Mobile Phone Image	Zhengqin Li*, UC San Diego; Manmohan Chandraker, UC San Diego; Sunkavalli Kalyan, Adobe Research

1C - Video	September 10, 02:15 pm – 04:00 pm **Chairs:** Ivan Laptev, INRIA Thomas Brox, University of Freiburg	
	Title	**Authors**
O-1C-01	End-to-End Joint Semantic Segmentation of Actors and Actions in Video	Jingwei Ji*, Stanford University; Shyamal Buch, Stanford University; Alvaro Soto, Universidad Catolica de Chile; Juan Carlos Niebles, Stanford University
O-1C-02	Learning-based Video Motion Magnification	Tae-Hyun Oh, MIT CSAIL; Ronnachai Jaroensri*, MIT CSAIL; Changil Kim, MIT CSAIL; Mohamed A. Elghareb, Qatar Computing Research Institute; Fredo Durand, MIT; Bill Freeman, MIT; Wojciech Matusik, Adobe
O-1C-03	Massively Parallel Video Networks	Viorica Patraucean*, DeepMind; Joao Carreira, DeepMind; Laurent Mazare, DeepMind; Simon Osindero, DeepMind; Andrew Zisserman, University of Oxford
O-1C-04	DeepWrinkles: Accurate and Realistic Clothing Modeling	Zorah Laehner, TU Munich; Tony Tung*, Facebook / Oculus Research; Daniel Cremers, TUM
O-1C-05	Learning Discriminative Video Representations Using Adversarial Perturbations	Jue Wang*, ANU; Anoop Cherian, MERL

Poster Session 1B	September 10, 04:00 pm – 06:00pm, see page 153

Demo Session 1B	September 10, 04:00pm – 06:00pm, see page 100

2A - Humans analysis 1	September 11, 08:30 am – 09:45 am Chairs: Kris Kitani, Carnegie Mellon University Tinne Tuytelaars, KU Leuven	
	Title	Authors
O-2A-01	Scaling Egocentric Vision: The E-Kitchens Dataset	Dima Damen*, University of Bristol; Hazel Doughty, University of Bristol; Sanja Fidler, University of Toronto; Antonino Furnari, University of Catania; Evangelos Kazakos, University of Bristol; Giovanni Farinella, University of Catania, Italy; Davide Moltisanti, University of Bristol; Jonathan Munro, University of Bristol; Toby Perrett, University of Bristol; Will Price, University of Bristol; Michael Wray, University of Bristol
O-2A-02	Unsupervised Person Re-identification by Deep Learning Tracklet Association	Minxian Li*, Nanjing University and Science and Technology; Xiatian Zhu, Queen Mary University, London, UK; Shaogang Gong, Queen Mary University of London
O-2A-03	Predicting Gaze in Egocentric Video by Learning Task-dependent Attention Transition	Yifei Huang*, The University of Tokyo
O-2A-04	Instance-level Human Parsing via Part Grouping Network	Ke Gong*, SYSU; Xiaodan Liang, Carnegie Mellon University; Yicheng Li, Sun Yat-sen University; Yimin Chen, sensetime; Liang Lin, Sun Yat-sen University
O-2A-05	Adversarial Geometry-Aware Human Motion Prediction	Liangyan Gui*, Carnegie Mellon University; XIAODAN LIANG, Carnegie Mellon University; Yu-Xiong Wang, Carnegie Mellon University; José M. F. Moura, Carnegie Mellon University

Poster Session 2A	September 11, 10:00 am – 12:00 pm, see page 166

Demo Session 2A	September 11, 10:00 am – 12:00 pm, see page 103

Lunch	12:00 pm – 01:00 pm

2B – Human Sensing I	September 11, 01:00 pm – 02:15 pm **Chairs:** Mykhaylo Andriluka, Max Planck Insititute Pascal Fua, EPFL	
	Title	Authors
O-2B-01	Weakly-supervised 3D Hand Pose Estimation from Monocular RGB Images	Yujun Cai*, Nanyang Technological University; Liuhao Ge, NTU; Jianfei Cai, Nanyang Technological University; Junsong Yuan, State University of New York at Buffalo, USA
O-2B-02	Audio-Visual Scene Analysis with Self-Supervised Multisensory Features	Andrew Owens*, UC Berkeley; Alexei Efros, UC Berkeley
O-2B-03	Jointly Discovering Visual Objects and Spoken Words from Raw Sensory Input	David Harwath*, MIT CSAIL; Adria Recasens, Massachusetts Institute of Technology; Dídac Surís, Universitat Politecnica de Catalunya; Galen Chuang, MIT; Antonio Torralba, MIT; James Glass, MIT
O-2B-04	DeepIM: Deep Iterative Matching for 6D Pose Estimation	Yi Li*, Tsinghua University; Gu Wang, Tsinghua University; Xiangyang Ji, Tsinghua University; Yu Xiang, University of Michigan; Dieter Fox, University of Washington
O-2B-05	Implicit 3D Orientation Learning for 6D Object Detection from RGB Images	Martin Sundermeyer*, German Aerospace Center (DLR); Zoltan Marton, DLR; Maximilian Durner, DLR; Rudolph Triebel, German Aerospace Center (DLR)

2C – Computational Photograpy 2	September 11, 02:45 pm – 04:00 pm Chairs: Kyros Kutulakos, University of Toronto Kalyan Sunkavalli, Adobe Research	
	Title	Authors
O-2C-01	Direct Sparse Odometry With Rolling Shutter	David Schubert*, Technical University of Munich; Vladyslav Usenko, TU Munich; Nikolaus Demmel, TUM; Joerg Stueckler, Technical University of Munich; Daniel Cremers, TUM
O-2C-02	3D Motion Sensing from 4D Light Field Gradients	Sizhuo Ma*, University of Wisconsin-Madison; Brandon Smith, University of Wisconsin-Madison; Mohit Gupta, University of Wisconsin-Madison, USA
O-2C-03	A Style-aware Content Loss for Real-time HD Style Transfer	Artsiom Sanakoyeu*, Heidelberg University; Dmytro Kotovenko, Heidelberg University; Bjorn Ommer, Heidelberg University
O-2C-04	Scale-Awareness of Light Field Camera based Visual Odometry	Niclas Zeller*, Karlsruhe University of Applied Sciences; Franz Quint, Karlsruhe University of Applied Sciences; Uwe Stilla, Technische Universitaet Muenchen
O-2C-05	Burst Image Deblurring Using Permutation Invariant Convolutional Neural Networks	Miika Aittala*, MIT; Fredo Durand, MIT

Poster Session 2B	September 11, 04:00 pm – 06:00 pm, see page 178

Demo Session 2B	September 11, 04:00 pm – 06:00 pm, see page 106

3A – Stereo and reconstruction	September 12, 08:30 am – 09:45 am Chairs: Noah Snavely, Cornell University Andreas Geiger, University of Tübingen	
	Title	Authors
O-3A-01	MVSNet: Depth Inference for Unstructured Multi-view Stereo	Yao Yao*, The Hong Kong University of Science and Technology; Zixin Luo, HKUST; Shiwei Li, HKUST; Tian Fang, HKUST; Long Quan, Hong Kong University of Science and Technology
O-3A-02	PlaneMatch: Patch Coplanarity Prediction for Robust RGB-D Registration	Yifei Shi, Princeton University; Kai Xu, Princeton University and National University of Defense Technology; Matthias Niessner, Technical University of Munich; Szymon Rusinkiewicz, Princeton University; Thomas Funkhouser*, Princeton, USA
O-3A-03	Active Stereo Net: End-to-End Self-Supervised Learning for Active Stereo Systems	Yinda Zhang*, Princeton University; Sean Fanello, Google; Sameh Khamis, Google; Christoph Rhemann, Google; Julien Valentin, Google; Adarsh Kowdle, Google; Vladimir Tankovich, Google; Shahram Izadi, Google; Thomas Funkhouser, Princeton, USA
O-3A-04	GAL: Geometric Adversarial Loss for Single-View 3D-Object Reconstruction	Li Jiang*, The Chinese University of Hong Kong; Xiaojuan Qi, CUHK; Shaoshuai SHI, The Chinese University of Hong Kong; Jia Jiaya, Chinese University of Hong Kong
O-3A-05	Deep Virtual Stereo Odometry: Leveraging Deep Depth Prediction for Monocular Direct Sparse Odometry	Nan Yang*, Technical University of Munich; Rui Wang, Technical University of Munich; Joerg Stueckler, Technical University of Munich; Daniel Cremers, TUM

Poster Session 3A	September 12, 10:00 am – 12:00 pm, see page 190

Demo Session 3A	September 12, 10:00 am – 12:00 pm, see page 110

Lunch	12:00 pm – 01:00 pm

3B - Human Sensing II	September 12, 01:00 pm – 02:15 pm **Chairs:** Gerard-Pons Moll, Max Planck Institute Juergen Gall, University of Bonn	
	Title	Authors
O-3B-01	Unsupervised Geometry-Aware Representation for 3D Human Pose Estimation	Helge Rhodin*, EPFL; Mathieu Salzmann, EPFL; Pascal Fua, EPFL, Switzerland
O-3B-02	Dual-Agent Deep Reinforcement Learning for Deformable Face Tracking	Minghao Guo, Tsinghua University; Jiwen Lu*, Tsinghua University; Jie Zhou, Tsinghua University, China
O-3B-03	Deep Autoencoder for Combined Human Pose Estimation and Body Model Upscaling	Matthew Trumble*, University of Surrey; Andrew Gilbert, University of Surrey; John Collomosse, Adobe Research; Adrian Hilton, University of Surrey
O-3B-04	Occlusion-aware Hand Pose Estimation Using Hierarchical Mixture Density Network	Qi Ye*, Imperial College London; Tae-Kyun Kim, Imperial College London
O-3B-05	GANimation: Anatomically-aware Facial Animation from a Single Image	Albert Pumarola*, Institut de Robotica i Informatica Industrial; Antonio Agudo, Institut de Robotica i Informatica Industrial, CSIC-UPC; Aleix Martinez, The Ohio State University; Alberto Sanfeliu, Industrial Robotics Institute; Francesc Moreno, IRI

Poster Session 3B	September 12, 02:30 pm – 04:00 pm, see page 203

Demo Session 3B	September 12, 02:30 pm – 04:00 pm, see page 113

3C - Optimization	September 12, 04:00 pm – 05:15 pm Chairs: Vincent Lepetit, University of Bordeaux Vladlen Koltun, Intel	
	Title	Authors
O-3C-01	Deterministic Consensus Maximization with Biconvex Programming	Zhipeng Cai*, The University of Adelaide; Tat-Jun Chin, University of Adelaide; Huu Le, University of Adelaide; David Suter, University of Adelaide
O-3C-02	Robust fitting in computer vision: easy or hard?	Tat-Jun Chin*, University of Adelaide; Zhipeng Cai, The University of Adelaide; Frank Neumann, The University of Adelaide, School of Computer Science, Faculty of Engineering, Computer and Mathematical Science
O-3C-03	Highly-Economized Multi-View Binary Compression for Scalable Image Clustering	Zheng Zhang*, Harbin Institute of Technology Shenzhen Graduate School; Li Liu, the inception institute of artificial intelligence; Jie Qin, ETH Zurich; Fan Zhu, the inception institute of artificial intelligence ; Fumin Shen, UESTC; Yong Xu, Harbin Institute of Technology Shenzhen Graduate School; Ling Shao, Inception Institute of Artificial Intelligence; Heng Tao Shen, University of Electronic Science and Technology of China (UESTC)
O-3C-04	Efficient Semantic Scene Completion Network with Spatial Group Convolution	Jiahui Zhang*, Tsinghua University; Hao Zhao, Intel Labs China; Anbang Yao, Intel Labs China; Yurong Chen, Intel Labs China; Hongen Liao, Tsinghua University

O-3C-05	Asynchronous, Photometric Feature Tracking using Events and Frames	Daniel Gehrig, University of Zurich; Henri Rebecq*, University of Zurich; Guillermo Gallego, University of Zurich; Davide Scaramuzza, University of Zurich& ETH Zurich, Switzerland

Poster Session 3C	September 12, 05:15 pm – 06:45 pm, see page 213

Demo Session 3B	September 12, 5:15 pm – 06:00 pm, see page 113

Main Conference – Thursday, 13 September 2018

4A - Learning for Vision 2	September 13, 08:30 am – 09:30 am **Chairs:** Kyoung Mu Lee, Seoul National University Michael Felsberg, Linköping University	
	Title	**Authors**
O-4A-01	Group Normalization	Yuxin Wu, Facebook; Kaiming He*, Facebook Inc., USA
O-4A-02	Deep Expander Networks: Efficient Deep Networks from Graph Theory	Ameya Prabhu*, IIIT Hyderabad; Girish Varma, IIIT Hyderabad; Anoop Namboodiri, IIIT Hyderbad
O-4A-03	Towards Realistic Predictors	Pei Wang*, UC San Diego; Nuno Vasconcelos, UC San Diego
O-4A-04	Learning SO(3) Equivariant Representations with Spherical CNNs	Carlos Esteves*, University of Pennsylvania; Kostas Daniilidis, University of Pennsylvania; Ameesh Makadia, Google Research; Christine Allec-Blanchette, University of Pennsylvania

Poster Session 4A	September 13, 10:00 am – 12:00 pm, see page 223

Demo Session 4A	September 13, 10:00 am – 12:00 pm, see page 115

Lunch	12:00 pm – 01:00 pm

4B - Matching and Recognition	September 13, 01:00 pm – 02:15 pm Chairs: Ross Girshick, Facebook Philipp Kraehenbuehl, University of Texas at Austin	
	Title	Authors
O-4B-01	CornerNet: Detecting Objects as Paired Keypoints	Hei Law*, University of Michigan; Jia Deng, University of Michigan
O-4B-02	RelocNet: Continous Metric Learning Relocalisation using Neural Nets	Vassileios Balntas*, University of Oxford; Victor Prisacariu, University of Oxford; Shuda Li, University of Oxford
O-4B-03	The Contextual Loss for Image Transformation with Non-Aligned Data	Roey Mechrez*, Technion; Itamar Talmi, Technion; Lihi Zelnik-Manor, Technion
O-4B-04	Acquisition of Localization Confidence for Accurate Object Detection	Borui Jiang*, Peking University; Ruixuan Luo, Peking University; Jiayuan Mao, Tsinghua University; Tete Xiao, Peking University; Yuning Jiang, Megvii(Face++) Inc
O-4B-05	Deep Model-Based 6D Pose Refinement in RGB	Fabian Manhardt*, TU Munich; Wadim Kehl, Toyota Research Institute; Nassir Navab, Technische Universität München, Germany; Federico Tombari, Technical University of Munich, Germany

4C - Video and attention	September 13, 02:45 pm – 04:00 pm Chairs: Hedvig Kjellström, KTH Lihi Zelnik Manor, Technion	
	Title	Authors
O-4C-01	DeepTAM: Deep Tracking and Mapping	Huizhong Zhou*, University of Freiburg; Benjamin Ummenhofer, University of Freiburg; Thomas Brox, University of Freiburg
O-4C-02	ContextVP: Fully Context-Aware Video Prediction	Wonmin Byeon*, NVIDIA; Qin Wang, ETH Zurich; Rupesh Kumar Srivastava, NNAISENSE; Petros Koumoutsakos, ETH Zurich
O-4C-03	Saliency Benchmarking Made Easy: Separating Models, Maps and Metrics	Matthias Kümmerer*, University of Tübingen; Thomas Wallis, University of Tübingen; Matthias Bethge, University of Tübingen
O-4C-04	Museum Exhibit Identification Challenge for the Supervised Domain Adaptation.	Piotr Koniusz*, Data61/CSIRO, ANU; Yusuf Tas, Data61; Hongguang Zhang, Australian National University; Mehrtash Harandi, Monash University; Fatih Porikli, ANU; Rui Zhang, University of Canberra
O-4C-05	Multi-Attention Multi-Class Constraint for Fine-grained Image Recognition	Ming Sun, baidu; Yuchen Yuan, Baidu Inc.; Feng Zhou*, Baidu Research; Errui Ding, Baidu Inc.

Poster Session 4B	September 13, 04:00pm – 06:00pm, see page 235

Demo Session 4B	September 13, 04:00pm – 06:00pm, see page 119

Station	Demo session 1A	September 10, 10:00 am - 12:00 pm
1	Lip Movements Generation at a Glance	Lele Chen, Zhiheng Li, Sefik Emre Eskimez, Ross Maddox, Zhiyao Duan, Chenliang Xu (University of Rochester)

Cross-modality generation is an emerging topic that aims to synthesize data in one modality based on information in a different modality. In this work we consider a task of such: given an arbitrary audio speech and one face image of arbitrary target identity, generate synthesized facial movements of the target identity saying the speech. To perform well in this task, it inevitably requires a model to not only consider the retention of target identity, photo-realistic of synthesized images, consistency and smoothness of face images in a sequence, but more importantly, learn the correlations between audio speech and lip movements. To solve the collective problems, we explore the best modeling of the audio-visual correlations in building and training a lip-movement generator network. Specifically, we devise a method to fuse audio and image embeddings to generate multiple lip images at once and propose a novel method to synchronize lip changes and speech changes. Our model is trained in an end-to-end fashion and is robust to view angles and different facial characteristics. Thoughtful experiments on different images ranging from male, female to cartoon character or even animal images to show that our model is robust and useful. Our model is trained on English speech and can test on other languages like German, Chinese and so on. For the demo video, please refer to https://youtu.be/mmI31GdGL5g.

2	CARLA: Democratizing Autonomous Driving Research	Felipe Codevilla (Computer Vision Center, Barcelona), Nestor Subiron (Computer Vision Center, Barcelona), Alexey Dosovitskiy (Intel Intelligent Systems Lab, Munich), German Ros (Intel Intelligent Systems Lab, Santa Clara), Antonio M. Lopez (Computer Vision Center, Barcelona), Vladlen Koltun (Intel Intelligent Systems Lab, Santa Clara)

CARLA is an open-source simulator for autonomous driving research. It follows a client-server architecture, where the server runs physics and sensor simulations and the clients run the AI drivers. Server and clients exchange sensorial data (images, point clouds), ground truth (depth, semantic segmentation, GPS, 3D bounding boxes), privileged information (e.g. traffic infractions, collisions), and vehicle commands/state for supporting the training and testing of AI drivers. It allows to specify the sensor suite, and environmental conditions such as weather, illumination, number of traffic participants, etc. It includes benchmarks for making possible to compare different AIs under the same conditions. In line with popular real-world datasets such as KITTI and Cityscapes, CARLA also allows to develop vision-based algorithms for 2D/3D object detection, depth estimation, semantic segmentation, Visual SLAM, tracking, etc.

CARLA was born for democratizing research on autonomous driving. Not only the source code is open and free to use/modify/redistribute, but also the 3D assets that are used to build the cities, i.e. buildings, roads, sidewalks, pedestrians, cars, bikes, motorbikes, etc. Since recently CARLA is member of the Open Source Vision Foundation (OSVF), so being the sister of OpenCV and Open3D.

In the seminar paper of CARLA (CoRL'2017), we studied the performance of three vision-based approaches to autonomous driving. Since its public release at during November 2017, many users have joint github CARLA community and have provided additional functionalities that were not originally released such, e.g. a LIDAR sensor and a ROS bridge. Recent interesting works use CARLA for proposing new vision-based approaches to autonomous driving, for instance:

A. Sauer, N. Savinov, A. Geiger, Conditional affordance learning for driving in urban environments, arXiv:1806.06498, 2018.
Video: https://www.youtube.com/watch?v=UtUbpigMgr0

Udacity Lyft perception challenge,
Video: https://www.youtube.com/watch?v=h17SnzLJAA4

CARLA premier video can be found at and more videos in: https://www.youtube.com/channel/UC1llP9ekCwt8nEJzMJBQekg

We will showcase CARLA in real-time, also showing videos of the most interesting papers up to date using CARLA to perform vision-based autonomous driving. Moreover, we will explain the development road-map of CARLA. Attached to the email there is a snapshot of CARLA environment running in real time. In the background there is the view of CARLA's server showing a vision-based AI driver controlling a car. In the foreground there is a view of a CARLA's client showing an on-board image with its depth and semantic segmentation, as well as a map placing the different active cars in the simulation.

3	Programmable Light Curtains	Jian Wang, Joseph Bartels, William Whittaker, Aswin Sankaranarayanan, Srinivasa Narasimhan (Carnegie Mellon University)

A vehicle on a road or a robot in the field does not need a full-blown 3D depth sensor to detect potential collisions or monitor its blind spot. Instead, it needs to only monitor if any object comes within its near proximity, which is an easier task than full depth scanning. We introduce a novel device that monitors the presence of objects on a virtual shell near the device, which we refer to as a light curtain. Light curtains offer a light-weight, resource-efficient and programmable approach for proximity awareness for obstacle avoidance and navigation. They also have additional benefits in terms of improving visibility in fog as well as flexibility in handling light fall-off. Our prototype for generating light curtains works by rapidly rotating a line sensor and a line laser, in synchrony. The device can generate light curtains of various shapes with a range of 20-30m in sunlight (40m under cloudy skies and 50m indoors) and adapts dynamically to the demands of the task. This interactive demo will showcase the potential of light curtains for applications such as safe-zone monitoring, depth imaging, and self-driving cars. This research was accepted for oral presentation at ECCV 2018.

4	Eyes of Things	O. Deniz, N. Vallez, J.L. Espinosa-Aranda, J.M. Rico, J. Parra (University of Castilla-La Mancha)

Eyes of Things (EoT) (www.eyesofthings.eu) is an Innovation Project funded by the European Commission within the Horizon 2020 Framework Programme for Research and Innovation. The objective in EoT has been to build an optimized core vision platform that can work independently and also embedded into all types of artefacts. The platform has been optimized for high-performance, low power-consumption, size, cost and programmability. EoT aims at being a flexible platform for OEMs to develop computer vision-based products and services in short time.

The main elements of the 7x5cm EoT device are the groundbreaking low-power Myriad 2 SoC, a small low-power camera, low power Wi-Fi connectivity and a micro-SD card. The platform includes libraries for general-purpose image processing and computer vision, QR code recognition, Python scripting language (besides C/C++ language), deep learning inference, video streaming, robot control, audio input and output, efficient wireless messaging, connectivity with cloud services like Google Cloud Vision API or Firebase Cloud Service, etc.

The functionality of the EoT device is demonstrated with some cool applications:

· The Next Generation Museum Guide demonstrator. In this demonstrator, the EoT device is inside a headset which automatically recognizes the painting the visitor is looking at and then provides information about the painting via audio. Video: https://www.youtube.com/watch?v=QR5LoKMd-Q8c

· The Smart Doll with Emotion Recognition demonstrator embeds an EoT device inside a doll's head. Facial emotion recognition has been implemented so that the doll can assess the children emotional display reacting accordingly through audio feedback. This demonstrator uses deep learning inference for facial emotion recognition. All processing is done on the EoT board, powered by a LiPo battery. Video: https://www.youtube.com/watch?v=v3YtUWWxiN0

· Flexible Mobile Camera: This demonstrator is actually a set of functionalities useful for surveillance and including additional functionality provided in the cloud using images captured by the EoT device. Video: https://www.youtube.com/watch?v=JXKmmEsww5Q. An incarnation of the previous demonstrator is the 'Litterbug' application, which aims to detect illegal littering, Video: https://www.youtube.com/watch?v=dR-v17YuOcg

Station	Demo session 1B	September 10, 04:00 pm - 06:00 pm
1	3D Object Tracking and Pose Estimation for Cinema Visual Effects	Bogdan Bugaev, Anton Kryshchenko, Roman Belov, Sergei Krivokhatskii (KeenTools)

3D object tracking is an essential part of cinema visual effects pipeline. It's used for different tasks including color correction, stereo conversion, object replacement, texturing, etc. Typical tracking conditions in this field are characterized by various motion patterns and complicated scenes. We present a demo of 3D object tracking software dedicated for visual effects in the cinema. The software exploits ideas presented in ECCV 2018 paper 'Combining 3D Model Contour Energy and Keypoints for Object Tracking'. The paper describes an approach for monocular model-based 3D object pose estimation. Preliminary object pose can be found using a keypoint-based technique. The initial pose can be refined via optimization of the contour energy function. The energy determines the degree of correspondence between the contour of the model projection and edges on the image. Contour energy optimization doesn't require a preliminary training that allows to integrate it in visual effects production pipeline easily. We use this method to improve tracking and simplify user-defined object positioning by automatic pose estimation. The approach was tested on numerous real-world projects and OPT public benchmark dataset.

2	Towards Real-time Learning of Monocular Depth Estimation Enabling Multiple View Synthesis on CPU	Matteo Poggi, Fabio Tosi, Stefano Mattoccia (University of Bologna)

The ability to infer depth from a single image in an unsupervised manner is highly desirable in several applications such as augmented reality, robotic, autonomous driving and so on. This topic represents a very challenging task, and the advent of deep learning enabled to tackle this problem with excellent results. Recently, we have shown in [1] how, by designing thin architectures, accurate monocular depth estimation can be carried out in real-time on devices with standard CPUs and even on low-powered devices (as reported in this video, https://www.youtube.com/watch?v=Q6ao-4JruIns). For instance, our model infers depth on a Raspberry Pi 3 at about 2 fps [1] with a negligible loss of accuracy compared to the state-of-the-art monocular method represented by Godard et al CVPR 2017.
Eventually, we have proposed in [2] a novel methodology to tackle unsupervised monocular depth estimation enabling to achieve better accuracy than the state-of-the-art. Such network also allows for synthesizing two novel views, never seen at training time, that can be used for interesting additional purposes. To prove the effectiveness of our network for this latter task, given the input image and the two novel synthesized views, we feed to a popular stereo algorithm (i.e., SGM in the attached video) different combinations of stereo pairs (in the video, synthesized left and input image, synthesized left and right) achieving consistent results.
The live demo will show how fast and accurate monocular depth estimation is feasible even on standard CPU-based architectures, including embedded devices. Moreover, our network [2] allows inferring novel synthesized views from the single input video stream. According to the reasons explained so far, we think that several application domains would benefit from these achievements and consequently attract the interest of ECCV 2018 attendants. We plan to organize a live demo using standard computing devices (notebooks, embedded devices, etc).

[1] M. Poggi, F. Aleotti, F. Tosi, S. Mattoccia, „Towards real-time unsupervised monocular depth estimation on CPU“, accepted at IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2018)
[2] M. Poggi, F. Tosi, S. Mattoccia, „Learning monocular depth estimation with unsupervised trinocular assumptions“, accepted at 6th International Conference on 3D Vision (3DV 2018)

3	Implicit 3D Orientation Learning for 6D Object Detection from RGB Images	Martin Sundermeyer, Zoltan-Csaba Marton, Maximilian Durner, Rudolph Triebel (German Aerospace Center (DLR))

In our Demo, we show real-time Object Detection and 6D Pose Estimation of multiple objects from single RGB images. The algorithm is built upon the main results of our paper which is presented during the oral session ("Implicit 3D Orientation Learning for 6D Object Detection from RGB Images", ID: 2662). The considered objects contain few texture and include (view-dependent) symmetries. These circumstances cause problems for many existing approaches. Our so-called 'Augmented Autoencoders' are solely trained on synthetic RGB data rendered from a 3D model using Domain Randomization. Thus, we do not require real pose-annotated image data and generalize to various test sensors and environments. Furthermore, we also run the demo on an embedded Nvidia Jetson TX2 to demonstrate the efficiency of our approach.

Station	Demo session 2A	September 11, 10:00 am - 12:00 pm
1	Practical Near-Field Photometric Stereo	Roberto Mecca, Fotis Logothetis, Roberto Cipolla (Italian Institute of Technology)

The proposed demo is a prototype of 3D scanner that uses photometric imaging in the near field for highly accurate shape reconstructions. It consists of a set of white-light LEDs synchronised with an RGB camera through a microcontroller. The 3D shape is retrieved by inferring 3D geometry from shading cues of an object lit by calibrated light sources.

The novelty of this prototype with respect to the state-of-the-art is the capability of working in the near-field. The advantage of having the inspected object close to the device is twofold. Firstly, very high spacial frequencies can be retrieved with the limit of precision being around 50 microns (by using a 8mm lens and a 3.2MP camera at around 4cm away from the object). Secondly, the proximity of the light sources to the object allows a higher signal-to-noise ratio with respect to the ambient lighting. This means that, differently from other photometric imaging based 3D scanners, our prototype can be used in an open environment.

The acquisition process consists of a number of LEDs (typically 8) flashing while the camera captures a RAW image per LED. The acquisition speed is ruled by the camera framerate. For the proposed demo, the acquisition task is achieved at ~40fps (that is ~25ms per LED).

The implementation of the 3D shape reconstruction runs on a laptop using an embedded GPU. The output consists of a .stl mesh having a number of vertices proportional to the number of pixels of the camera.

2	Activity-Preserving Face Anonymization for Privacy Protection	Zhongzheng Ren, Yong Jae Lee, Hyun Jong Yang, Michael S. Ryoo (EgoVid Inc.)

There is an increasing concern in computer vision devices invading the privacy of their users by recording unwanted videos. On one hand, we want the camera systems/robots to recognize important events and assist human daily life by understanding its videos, but on the other hand we also want to ensure that they do not intrude people's privacy. In this demo, we present a new principled approach for learning a video face anonymizer. We use an adversarial training setting in which two competing systems fight: (1) a video anonymizer that modifies the original video to remove privacy sensitive information (i.e., human face) while still trying to maximize spatial action detection performance, and (2) a discriminator that tries to extract privacy sensitive information from such anonymized videos. The end result is a video anonymizer that performs a pixel level modification to anonymize each person's face, with minimal effect on action detection performance. We experimentally confirm the benefit of our approach compared to conventional handcrafted video face anonymization methods including masking, blurring, and noise adding. See the project page https://jason718.github.io/project/privacy/main.html for a demo video and more results.

3	Inner Space Preserving Generative Pose Machine	Shuangjun Liu, Sarah Ostadabbas (Northeastern University)

Photographs are important because they seem to capture so much: in the right photograph we can almost feel the sunlight, smell the ocean breeze, and see the fluttering of the birds. And yet, none of this information is actually present in a two-dimensional image. Our human knowledge and prior experience allow us to recreate ``much'' of the world state (i.e. its inner space) and even fill in missing portions of occluded objects in an image since the manifold of probable world states has a lower dimension than the world state space. Like humans, deep networks can use context and learned ``knowledge'' to fill in missing elements. But more than that, if trained properly, they can modify (repose) a portion of the inner space while preserving the rest, allowing us to significantly change portions of the image. In this work, we present a novel deep learning based generative model that takes an image and pose specification and creates a similar image in which a target element is reposed.

In reposing a figure there are three goals: (a) the output image should look like a realistic image in the style of the source image, (b) the figure should be in the specified pose, and (c) the rest of the image should be as similar to the original as possible. Generative adversarial networks (GANs) are the "classic'' approach to solving the first goal by generating novel images that match a certain style. The second goal, putting the figure in the correct pose, requires a more controlled generation approach, such as conditional GANs (cGAN). At a superficial level, this seems to solve the reposing problem. However, these existing approaches generally either focus on preserving the image (goal c) or generating an entirely novel image based on the contextual image (goal b), but not both.

We address the problem of articulated figure reposing while preserving the image's inner space (goal b and c) via the introduction of our inner space preserving generative pose machine (ISP-GPM) that generates realistic reposed images (goal a). In ISP-GPM, an interpretable low-dimensional pose descriptor (LDPD) is assigned to the specified figure in the 2D image domain. Altering LDPD causes figure to be reposed. For image regeneration, we used stack of augmented hourglass networks in the cGAN framework, conditioned on both LDPD and the original image. Furthermore, we extended the ``pose'' concept to a more general format which is no longer a simple rotation of a single rigid body, and instead the relative relationship between all the physical entities captured in an image and also its background. A direct outcome of ISP-GPM is that by altering the pose state in an image, we can achieve unlimited generative reinterpretation of the original world, which ultimately leads to a one-shot data augmentation with the original image inner space preserved.

Station	Demo session 2B	September 11, 04:00 pm - 06:00 pm
1	Dynamic Multi-modal Segmentation on the Wild using a Scalable and Distributed Deep Learning Architecture	Edgar Andres Margffoy Tuay (Universidad de los Andes, Colombia)

A scalable, disponible and generic Deep Learning service architecture for mobile/web applications is presented. This architecture integrates with modern cloud services, such as Cloud Storage and Push Notifications via cloud messaging (Firebase), as well with known distributed technologies based on the Erlang programming language.

To demostrate such approach, we present a novel Android application to perform object instance segmentation based on natural language expressions. The application allows users to take a photo using the device camera or to pick a previous image present on the image gallery. Finally, we also present a novel web client that performs similar functions that is supported by the same architecture backbone and also integrates with state-of-the-art technologies and specifications such as ES6 Javascript and WebAssembly.

2	Single Image Water Hazard Detection using FCN for Autonomous Car	Xiaofeng Han, Chuong Nguyen, Shaodi You, Jianfeng Lu (CSIRO data61)

Water bodies, such as puddles and flooded areas, on and off road pose significant risks to autonomous cars. Detecting water from moving camera is a challenging task as water surface is highly reflective, and its appearance varies with viewing angle, surrounding scene, and weather conditions.

We will present a live demo (running on a GPU laptop) of our water puddle detection method based on a Fully Convolutional Network (FCN) with our newly proposed Reflection Attention Units (RAUs). An RAU is a deep network unit designed to embody the physics of reflection on water surface from sky and nearby scene. We show that FCN-8s with RAUs significantly improves precision and recall metrics as compared to FCN-8s, DeepLab V2 and Gaussian Mixture Model (GMM).

Links:
- Paper summary: https://youtu.be/Cwfc0HpuuOI
- Recorded demo on road: https://youtu.be/OUNk8yBdaMg
- Recorded demo off road: https://youtu.be/SHuuIq2IfEQ
- Codes: https://github.com/Cow911/SingleImageWaterHazardDetectionWithRAU.git

3	Real-time 3D Object Localisation on a Smartphone	Alessio Del Bue, Paul Gay, Yiming Wang, Stuart James (Italian Institute of Technology)

The demo will show a system implemented on a mobile phone that is able to automatically localise and display the position of 3D objects using RGB images only. With more details, the algorithm reconstruct the position and occupancy of rigid objects from a set of object detections in a video sequence and the respective camera poses captured by the smartphone camera. In practice, the user scans the environment by moving the device over the space and then he/she receives automatically object proposals from the system together with their localization in 3D. Technically, the algorithm first fits an ellipse onto the image plane at each bounding box as given by the object detector. We then infer the general 3D spatial occupancy and orientation of each object by estimating a quadric (ellipsoid) in 3D given the conics in different views. The use of a closed form solution offers a fast method which can be used in situ to construct a 3D representation of the recorded scene. The system allows also to label the objects with additional information like personalised notes, videos, images and html links. At the moment, the implementation is using a Tango phone but ongoing work is extending the system to ARCore and ARKit enabled smartphones.

The system can be used as a content generation tool for Augmented Reality as we can provide additional information anchored to physical objects. Alternatively, the system is employed as an annotation tool for creating datasets as we record different type of data (images, depth, objects position and camera trajectory in 3D). Ongoing research activities are embedding our system into various robotic platforms at the Italian Institute of Technology (iCub, Centaur, Coman, Walkman) in order to provide robots with object perception and reasoning capabilities.

4	Real-time Automatic Instrument Status Monitoring using Deep Neural Networks	Eunseop Lee, Junghyun Hong, Daijin Kim (POSTECH)

For smart factory, we need to monitor the status of analog instrumental sensors (idle/normal/danger state) automatically in real-time. To achieve this goal, we perform several tasks: (1) detect instrument, (2) detect numerals, (3) segment and recognize digits, (4) detect needle, and (5) compute needle value and decide the instrument status. Previous approaches used feature-based methods commonly, so they can detect and recognize instruments only for the predefined shapes with a low performance. To overcome these limitations, we propose the deep learning based approaches. First, the instrument detector network is similar with the existing YOLOv2 architecture, but we use only one anchor, not five anchor boxes as in the original YOLOv2, because the instruments commonly have two shapes such as circle and rectangle. To train the detection model, we prepare a plenty of instrument image dataset that includes a total of 18,000 training samples generated by data augmentation and 10,000 artificial samples generated from BEGAN. Second, the numerals detector network takes the Textboxes++ network which is the state-of-the-art text detector, which is based on SSD. So, we can detect numerals accurately with a high speed because Textboxes++ detects numerals precisely even on the poor conditions with quadrilateral bounding boxes. Third, the numeral recognition network consists of two functional blocks: (1) it separates the detected numeral region into several single digits using the maximally stable extremal regions (MSER) algorithm, and (2) it accepts each segmented digit to the CNN with five convolution layers and two fully connected layers. This network classifies each image patch into a 0~9 digit value. Then, we obtain a set of numbers corresponding to each numeral bounding box by regrouping the recognized digits. Fourth, we find the needle using the Edline edge detection algorithm and non-maximum suppression and we may get two major end-points from the needle. Finally, we calculate two angles between the needle line and the directional lines pointing to the smallest and the largest number. From these angles and two numbers, we estimate the needle value pointed by the needle Based on the estimated needle value, we can decide one of the instrument status. The proposed instrument monitoring system can detect instruments from the captured image by a camera and decide the current instrument status automatically in real-time. We get 93.8% accuracy over the test dataset for 15 fps on a Nvidia Titan Xp.

Station	Demo session 3A	September 12, 10:00 am - 12:00 pm
1	Computational Photography Software using Deep Learning: Perceptual Image Super-Resolution and Depth Image Refinement	Junho Jeon, Seong-Jin Park, Hyeongseok Son, Seungyong Lee (POSTECH)

In this demo, we introduce various computational photography software based on deep learning. These software have been developed for pursuing our ultimate goal of building a computational photography software library, which we call "COUPE." Currently COUPE offers several deep learning based image processing utilities, such as image aesthetic assessment, image composition assessment and enhancement, color transfer and enhancement, and non-blind deconvolution.

In the demo session, we will mainly show our recent developments on perceptual image super-resolution and depth image refinement, which will also be presented as posters at ECCV 2018. In addition, we will introduce a demo website where visitors can interactively run several computational photography software to obtain the results for their own images. A demo poster will provide more information on the software as well.

2	Uncertainty Estimates and Multi-Hypotheses Networks for Optical Flow	Eddy Ilg, Özgün Cicek, Silvio Galesso, Aaron Klein, Osama Makansi, Frank Hutter, Thomas Brox (University of Freiburg)

Recent work has shown that optical flow estimation can be formulated as an end-to-end supervised learning problem, which yields estimates with a superior accuracy-runtime trade-off compared to alternative methodology. However, for practical applications (e.g. autonomous driving), a major concern is how much the information can be trusted. In the past very little work has been done in this direction. In this work we show how state-of-the-art uncertainty estimation for optical flow can be obtained with CNNs. Our uncertainties generalize to real-world videos well, including challenges for optical flow such as self-occlusion, homogeneous areas and ambiguities. At the demo will show this live and in real-time.

3	Multi-Frame Quality Enhancement for Compressed Video	Ren Yang, Mai Xu, Zulin Wang, Tianyi Li (Beihang University)

The past few years have witnessed great success in applying deep learning to enhance the quality of compressed image/video. The existing approaches mainly focus on enhancing the quality of a single frame, ignoring the similarity between consecutive frames. This demo illustrates a novel Multi-Frame Quality Enhancement (MFQE) approach for compressed video, which was proposed in our CVPR'18 paper. In this work, we investigate that heavy quality fluctuation exists across compressed video frames, and thus low quality frames can be enhanced using the neighboring high quality frames, seen as Multi-Frame Quality Enhancement (MFQE). Accordingly, this work proposes an MFQE approach for compressed video, as a first attempt in this direction. In our approach, we firstly develop a Support Vector Machine (SVM) based detector to locate Peak Quality Frames (PQFs) in compressed video. Then, a novel Multi-Frame Convolutional Neural Network (MF-CNN) is designed to enhance the quality of compressed video, in which the non-PQF and its nearest two PQFs are as the input. The MF-CNN compensates motion between the non-PQF and PQFs through the Motion Compensation subnet (MC-subnet). Subsequently, the Quality Enhancement subnet (QE-subnet) reduces compression artifacts of the non-PQF with the help of its nearest PQFs. The experimental results validate the effectiveness and generality of our MFQE approach in advancing the state-of-the-art quality enhancement of compressed video.

4	Unsupervised Event-based Optical Flow Estimation using Motion Compensation	Alex Zihao Zhu, Liangzhe Yuan, Kenneth Chaney, Kostas Daniilidis (University of Pennsylvania)

This demo consists of a novel framework for unsupervised learning of optical flow for event cameras that learns from only the event stream. Event cameras are a novel sensing modality that asynchronously track changes in log light intensity. When a change is detected, the camera immediately sends an event, consisting of the x,y pixel position of the change, timestamp accurate to microseconds, and a polarity, indicating the direction of the change. The cameras provide a number of benefits over traditional cameras, such as low latency, the ability to track very fast motions, very high dynamic range, and low power consumption.

Our work in this demo allows us to show some of the capabilities of algorithms on this camera, by using a trained neural network to predict optical flow in very challenging environments from an event camera. Similar to EV-FlowNet, this work consists of a fully convolutional network that takes in a discretized representation of the event stream, and learns to predict optical flow for each event in a fully unsupervised manner. However, in this work, we propose a novel input representation consisting of a discretized 3D voxel grid, and a loss function that allows the network to learn optical flow from only the event stream by minimizing the motion blur in the scene (no grayscale frames needed).

Our network runs on a laptop grade NVIDIA GTX 960M at 20Hz, and is able to track very fast motions such as objects spinning at 40rad/s, as well as in very challenging and varying lighting conditions, all in realtime. The network, trained only on 10 minutes of driving sequences, generalizes to a variety of different scenes, without any noticeable outliers. We encourage audience participation.

Station	Demo session 3B	September 12, 02:30 pm – 04:00 pm, 05:15 pm – 06:00 pm
1	Annotation Tool for Large CG Datasets	Matteo Fabbri, Fabio Lanzi, Andrea Palazzi, Roberto Vezzani, Rita Cucchiara (University of Modena and Reggio Emilia)

We present a demo of our automatic annotation tool used for the creation of the "JTA" Dataset (Fabbri et al. "Learning to Detect and Track Visible and Occluded Body Joints in a Virtual World". ECCV 2018). This tool allows you to quickly create tracking and pose detection datasets through a convenient graphical interface exploiting the highly photorealistic video game "GTA V" (see video: https://youtu.be/9Q1UYzUysUk).
During the demo we will demonstrate the ease with which our mod allows you to create new scenarios and control the behavior/number/type/appearance/interactions of people on the screen, showing at the same time the quality of the obtained annotations (both in terms of tracking, 2D and 3D pose detection). Currently, multi-person tracking and pose detection video datasets are small, as the manual annotation for these tasks is extremely complex and time-consuming; moreover the manual approach often does not guarantee optimal results due to unavoidable human errors and difficulties in reconstructing the correct poses of strongly occluded people. The code of our tool will be released soon.

2	Real-time Multi-person 3D Human Pose Estimation System	Dushyant Mehta, Oleksandr Sotnychenko, Franziska Mueller, Helge Rhodin, Weipeng Xu, Gerard Pons-Moll, Christian Theobalt (Max-Planck-Institute for Informatics)

We present a real-time multi-person 3D human body pose estimation system which makes use of a single RGB camera for human motion capture in general scenes. Our learning based approach gives full body 3D articulation estimates even under strong partial occlusion, as well as estimates of camera relative localization in space. Our approach makes use of the detected 2D body joint locations as well as the joint detection confidence values, and is trained using our recently proposed Multi-person Composited 3D Human Pose (MuCo-3DHP) dataset, and also leverages MS-COCO person keypoints dataset for improved performance in general scenes. Our system can handle an arbitrary number of people in the scene, and processes complete frames without requiring prior person detection.

For details, please refer to: http://people.mpi-inf.mpg.de/~dmehta/XNect-Demo/

3	LCR-Net++: Multi-person 2D and 3D Pose Detection in Natural Images	Gregory Rogez (Inria), Philippe Weinzaepfel (Naver Labs), Xavier Martin (Inria), Cordelia Schmid (Inria)

We present a real-time demo of our 3D human pose detector, called LCR-Net++, recently accepted to appear in IEEE TPAMI. This is an extended version of the LCR-Net paper published at CVPR'17. It is the first detector that always estimates the full-body 2D and 3D poses of multiple people from a single RGB image. Our approach significantly outperforms the state of the art in 3D pose estimation on Human3.6M, the standard benchmark dataset captured in a controlled environment, and demonstrates satisfying 3D pose results in real-world images, hallucinating plausible body parts when the persons are partially occluded or truncated by the image boundary.

More details on the project website: https://thoth.inrialpes.fr/src/LCR-Net/

4	HybridFusion: Real-Time Performance Capture Using a Single Depth Sensor and Sparse IMUs	Zerong Zheng (Tsinghua University), Tao Yu (Beihang University & Tsinghua University), Yebin Liu (Tsinghua University)

We propose a light-weight and highly robust real-time human performance capture method based on a single depth camera and sparse inertial measurement units (IMUs). The proposed method combines non-rigid surface tracking and volumetric surface fusion to simultaneously reconstruct challenging motions, detailed geometries and the inner human body shapes of a clothed subject. The proposed hybrid motion tracking algorithm and efficient per-frame sensor calibration technique enable non-rigid surface reconstruction for fast motions and challenging poses with severe occlusions. Significant fusion artifacts are reduced using a new confidence measurement for our adaptive TSDF-based fusion. The above contributions benefit each other in our real-time reconstruction system, which enable practical human performance capture that is real-time, robust, low-cost and easy to deploy. Our experiments show how extremely challenging performances and loop closure problems can be handled successfully.

Station	Demo session 4A	September 13, 10:00 am – 12:00 pm
1	Scalabel: A Versatile and Scalable Annotation System	Fisher Yu, Xin Wang, Sean Foley, Haofeng Chen, Haoping Bai, Gary Chen, Yingying Chen, Simon Mo, Ryan Cheng, Joseph Gonzalez, Trevor Darrell (UC Berkeley)

The success of deep learning depends on large-scale annotated vision dataset. However, there has been surprisingly little work on the design of tools for efficient data labeling at scale. To foster the systematic study of large-scale data annotation, we introduce Scalabel, a versatile and scalable annotation system that accommodates a wide range annotation tasks needed for the computer vision community and enables fast labeling with minimal expertise. We support both generic 2D/3D bounding boxes annotation, semantic/instance segmentation annotation, video object tracking as well as domain-specific annotation tasks for autonomous driving, such as lane detection, drivable area segmentation, etc. By providing a common API and visual language for data annotation across a wide range of annotation tasks, Scalabel is both easy to use and to extend. Our labeling system was used to annotate BDD100K dataset and it received positive feedback from the workers during this large-scale production. To our knowledge, existing open-sourced annotation tools are built for specific annotation tasks (e.g., single object detection or vehicle/pedestrian detection) that can not be readily extended to new tasks.

2	ARCADE - Accurate and Rapid Camera for Depth Estimation	Harel Haim, Shay Elmalem, Raja Giryes, Alex Bronstein and Emanuel Marom (Tel-Aviv University)

A depth estimation solution based on a single-shot taken with a single phase-coded aperture camera is proposed and presented. One of the most challenging tasks in computer vision is depth estimation from a single image. The main difficulty lies in the fact that depth information is lost in conventional 2D imaging. Various computational imaging approaches have been proposed to address this challenge, such as incorporating an amplitude-mask in the imaging system pupil. Such mask encodes subtle depth dependent cues in the resultant image, which are then used for depth estimation in the post-processing step. Yet, a proper design of such an element is challenging. Moreover, amplitude phase masks reduce the light throughput of the imaging system (in some cases up to 50%). The recent and ongoing Deep Learning (DL) revolution did not pass over this challenge, and many monocular depth estimation Convolutional Neural Networks (CNN) have been proposed. In such solutions, a CNN model is trained to estimate the scene depth map using labeled data. These methods rely on monocular depth cues (perspective, vanishing lines, proportion, etc.), which are not always clear and helpful. In this work, the DL end-to-end design approach is employed to jointly design a phase-mask and a DL model, working in tandem to achieve scene depth estimation from a single image. Utilizing phase aperture-coding has the important advantage of nearly 100% light throughput. The phase coded aperture imaging is modeled as a layer in the DL structure, and its parameters are jointly learned with the CNN to achieve true synergy between the optics and the post processing step. The designed phase mask encodes color and depth dependent cues in the image, which enable depth estimation using a relatively shallow FCN model, trained for depth reconstruction. After the training stage is completed, the phase element is fabricated (according to the trained optical parameters) and incorporated in the aperture stop of a conventional lens, mounted on a conventional camera. The Raw images taken with this phase-coded camera are fed to the 'conventional' FCN model, and depth estimation is achieved. In addition, utilizing the coded PSFs, an all-in focus image can be restored. Combining the all-in-focus image with the acquired depth map, synthetic re-focusing can be created, with the proper Bokeh effect.

3	Learn, Generate, Rank: Generative Ranking of Motion Capture	Xiao Lin, Chris Kim, Timothy Meo, and Mohamed R. Amer (SRI International)

We present a novel approach for searching and ranking videos for activities using deep generative model. Ranking is a well-established problem in computer vision. It is usually addressed using discriminative models, however, the decisions made by these models tend to be unexplainable. We believe that generative models are more explainable since they can generate instances of what they learned.

Our model is based on Generative Adversarial Networks (GANs). We formulate a Dense Validation GANs (DV-GANs) that learns human motion, generate realistic visual instances given textual inputs, and then uses the generated instances to search and rank videos in a database under a perceptually sound distance metric in video space. The distance metric can be chosen from a spectrum of handcrafted to learned distance functions controlling trade-offs between explainability and performance. Our model is capable of human motion generation and completion.

We formulate the GAN discriminator using a Convolutional Neural Network (CNN) with dense validation at each time-scale and perturb the discriminator input to make it translation invariant. Our DVGAN generator is capable of motion generation and completion using a Recurrent Neural Network (RNN). For encoding the textual query, a pretrained language models such as skip-thought vectors are used to improve robustness to unseen query words.

We evaluate our approach on Human 3.6M and CMU motion capture datasets using inception scores. Our approach shows through our evaluations the resiliency to noise, generalization over actions, and generation of long diverse sequences.

Our demo is available at: http://visxai.ckprototype.com/demo/ and A simplified blog post of the demo is available at: http://genrank.ckprototype.com/

4	Iterative Query Refinement with Visual Explanations	Bo Dong (Kitware), Vitali Petsiuk (Boston University), Abir Das (Boston University), Kate Saenko (Boston University), Roddy Collins (Kitware), Anthony Hoogs (Kitware)

Large-scale content-based image retrieval (CBIR) systems are widespread [1,2,3], yet most CBIR systems for reverse image search suffer from two drawbacks: inability to incorporate user feedback, and uncertainty why any particular image was retrieved. Without feedback, the user cannot communicate to the system which results are relevant and which are not; without explanations, the system cannot communicate to the user why it believes its answers are correct. Our demonstration proposes solutions to both problems, implemented in an open-source image query and retrieval framework. We demonstrate incorporation of user feedback to increase query precision, combined with a method to generate saliency maps which explain why the system believes that the retrievals match the query. We incorporate feedback via iterative query refinement (IQR), in which the user, via a web-based GUI, provides binary relevance feedback (positive or negative) to refine previously retrieved results until the desired precision is achieved. In each iteration, the feedback is used to train a two-class SVM classifier; this reranks the initial result set attempting to increase the likelihood that higher-ranked results would also garner positive feedback. Secondly, we generate saliency maps for each result, reflecting how regions in a retrieved image match the query image. Building on [4,5,6], we repeatedly obscure regions in the matched image and recompute the similarity metric to identify which regions most affect the score. Unlike previous methods, ours does not require matches to belong to predefined classes. These saliency maps provide the explanation, visualizing the underlying matching criteria to show why a retrieved image was matched. Generating saliency no additional models or modifications to the underlying algorithm. This is combined research of Kitware (IQR and framework) and Boston University (saliency maps), developed as part of the DARPA Explainable AI project.

A video of the system is at https://youtu.be/WTnWofS1nEE.

[1] http://tineye.com
[2] http://www.visenze.com
[3] http://www.snapfashion.co.uk
[4] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, “Learning deep features for discriminative localization,” CVPR 2016
[5] R. R. Selvaraju, A. Das, R. Vedantam, M. Cogswell, D. Parikh, and D. Batra, “Grad-cam: Why did you say that? visual explanations from deep networks via gradient-based localization,” arXiv preprint arXiv:1610.02391, 2016.
[6] A. Chattopadhyay, A. Sarkar, P. Howlader, and V. N. Balasubramanian. Grad-cam++: Generalized gradient-based visual explanations for deep convolutional networks. CoRR, abs/1710.11063, 2017.

Station	Demo session 4B	September 13, 04:00 pm – 06:00 pm
1	Camera Tracking for SLAM in Deformable Maps	Jose Lamarca, JMM Montiel (Universidad de Zaragoza)

The current VSLAM algorithms cannot work without assuming rigidity. We propose the first real-time tracking thread for VSLAM systems that manages deformable scenes. It is based on top of the Shape-from-Template (SfT) methods to code the scene deformation model. Our proposal is a sequential two-step method that manages efficiently large templates locating the camera at the same time. We show the system with a demo in which we move the camera just imaging a small part of a fabric while we deform it, recovering both deformation and camera pose in real-time (20Hz).

2	Ultimate SLAM: Combining Events, Frames and IMU for Robust Visual SLAM in HDR and High Speed Scenarios	Antoni Rosinol Vidal, Henri Rebecq, Timo Horstschaefer, Davide Scaramuzza (ETH Zurich and University of Zurich)

We would like to present a demo our paper [1], http://rpg.ifi.uzh.ch/ultimateslam.html
Event cameras are bio-inspired vision sensors that output pixel-level brightness changes instead of standard intensity frames. These cameras provide reliable visual information during high-speed motions or in scenes with high dynamic range; however, they output little information during slow motions. Conversely, standard cameras provide rich visual information most of the time (in low-speed and good lighting scenarios). We present the first SLAM pipeline that leverages the complementary advantages of these two sensors by fusing in a tightly-coupled manner events, intensity frames, and inertial measurements. We show that our pipeline leads to an accuracy improvement of 130% over event-only pipelines, and 85% over standard-frames-only visual-inertial systems, while still being computationally tractable.
We believe that event cameras are of great interest to ECCV audience, bringing exciting new ideas about asynchronous and sparse acquisition and processing of visual information. Event cameras are an emerging technology, supported by companies with multi-million investment, such as Samsung and Prohesee [2].
[1] A. Rosinol Vidal et al., Ultimate SLAM?, Combining Events, Images and IMU for Robust Visual SLAM in HDR and High Speed Scenarios, IEEE RA-L, 2018. http://rpg.ifi.uzh.ch/ultimateslam.html
[2] http://www.prophesee.ai/2018/02/21/prophesee-19-million-funding-round/

3	FlashFusion: Real-time Globally Consistent Dense 3D Reconstruction	Lei Han, Lu Fang (Tsinghua University)

Aiming at the practical usage of dense 3D reconstruction on portable devices, we propose FlashFusion, a Fast LArge-Scale High-resolution (sub-centimeter level) 3D reconstruction system without the use of GPU computing. It enables globally-consistent localization through a robust yet fast global bundle adjustment scheme, and realizes spatial hashing based volumetric fusion running at 300Hz and rendering at 25Hz via highly efficient valid chunk selection and mesh extraction schemes. Extensive experiments on both real world and synthetic datasets demonstrate that FlashFusion succeeds to enable real- time, globally consistent, high-resolution (5mm), and large-scale dense 3D reconstruction using highly-constrained computation, i.e., the CPU computing on portable device. The associate paper is previously published on Robotics, Science and Systems 2018 as oral presentation.

4	EVO: A Geometric Approach to Event-based 6-DOF Parallel Tracking and Mapping in Real-time	Henri Rebecq, Timo Horstschaefer, Guillermo Gallego, Davide Scaramuzza (ETH Zurich and University of Zurich)

We wish to present a live demo of our paper EVO [1], https://youtu.be/bYqD2qZJlxE. Event cameras are bio-inspired vision sensors that output pixel-level brightness changes instead of standard intensity frames. They offer significant advantages over standard cameras (a very high dynamic range, no motion blur, and microsecond latency). However, traditional vision algorithms cannot be applied to the output of these sensors, so that a paradigm shift is needed. Our structure from motion algorithm successfully leverages the outstanding properties of event cameras to track fast camera motions while recovering a semi-dense 3D reconstruction of the environment. Our work makes significant progress in SLAM by unlocking the potential of event cameras, allowing us to tackle challenging scenarios (e.g., high-speed) that are currently naccessible to standard cameras.
To the best of our knowledge, this is the first work showing real-time structure from motion on a CPU for an event camera moving in six degrees-of-freedom. We believe the paradigm shift posed by event cameras is of great interest to ECCV audience, bringing exciting new ideas about asynchronous and sparse acquisition (and processing) of visual information. Event cameras are an emerging technology that is attracting the attention of investment funding, such as the $40 Million raised by Prophesee [2] or the multi-million dollar investment by Samsung [3].

[1] Rebecq, Horstschaefer, Gallego and Sacaramuzza, A Geometric Approach to Event-based 6-DOF Parallel Tracking and Mapping in Real-time, IEEE RAL, 2017.
[2] http://www.prophesee.ai/2018/02/21/prophesee-19-million-funding-round/
[3] Samsung turns IBM's brain-like chip into a digital eye, https://www.cnet.com/news/samsung-turns-ibms-brain-like-chip-into-a-digital-eye/

Workshop Number: 3

The Visual Object Tracking Challenge Workshop VOT2018

Date: Friday 14th, morning

Room: N1179

Organizers: Matej Kristan, Aleš Leonardis, Jiři Matas, Michael Felsberg, Roman Pflugfelder

SCHEDULE

8:30 VOT Welcome

8:35 – 10:00 Oral session 1
- VOT Results + winners announcement
- Poster spotlights

10:00 – 11:00 **Poster session** (Coffee break during poster session)

- How to Make an RGBD Tracker?
Ugur Kart (Tampere University of Technology)*
Joni-Kristian Kamarainen (Tampere University of Technology , Finland); Jiri Matas (CMP CTU FEE)

- A Memory Model based on the Siamese Network for Long-term Tracking
Hankyeol Lee (KAIST)*; Seokeon Choi (KAIST); Changick Kim (KAIST)

- On the Optimization of Advanced DCF-Trackers
Joakim Johnander (Linköping University)*; Goutam Bhat (Linkoping University); Martin Danelljan (Linkoping University); Fahad Shahbaz Khan (Linköping University); Michael Felsberg (Linköping University)

- Channel Pruning for Visual Tracking
Manqiang Che (North China University of Technology); Runling Wang (North China University of Technology); Yan Lu (North China University of Technology); Yan Li (North China University of Technology); Hui Zhi (North China University of Technology); Changzhen Xiong (North China University of Technology)*

· Towards a Better Match in Siamese Network Based Visual Object Tracker
Anfeng He (USTC); Chong Luo (MSRA)*; Xinmei Tian (USTC); Wenjun Zeng (Microsoft Research)

· Learning a Robust Society of Tracking Parts using Co-occurrence Constraints
Elena Burceanu (Bitdefender, Institute of Mathematics of the Romanian Academy)*; Marius Leordeanu (Institute of Mathematics of the Romanian Academy)

·. WAEF: Weighted Aggregation with Enhancement Filter for Visual Object Tracking
Litu Rout (Indian Institute of Space Science and Technology)*; Deepak Mishra (IIST); Gorthi Rama Krishna Sai Subrahmanyam (IIT Tirupati)

· Multiple Context Features in Siamese Networks for Visual Object Tracking
Henrique Morimitsu (Inria)*

11:00 – 12:35 **Oral session 2**

· Invited talk: J.-C. Olivo-Marin
(Institut Pasteur, BioImage Analysis Unit)
· VOT Short-term challenge second-best tracker talk
· VOT Real-time challenge winners talk
· VOT Long-term challenge winners talk
· Panel discussion & concluding remarks

Workshop Number: 14

The Third International Workshop on Video Segmentation

Date: Friday 14th, morning

Room: 1200 Carl Von Linde Hall

Organizers: Pablo Arbelaez, Thomas Brox, Fabio Galasso, Iasonas Kokkinos, Fuxin Li

SCHEDULE

8:40	Introduction
8:45	Invited speaker: Cordelia Schmid. „Learning to segment moving objects“
9:15	Invited short talk: Jordi Pont-Tuset. „Two years of DAVIS: what we've learnt and what's next in video object segmentation“
9:30	Invited speaker: Daniel Cremers. “Inverse Problems in the Age of Deep Learning”
10:00	Khoreva et al. „Video Object Segmentation with Referring Expressions“
10:15	Coffee break
10:45	Invited speaker: Bernt Schiele. „Video segmentation with less supervision“
11:15	Jain et al. „Fast Semantic Segmentation on Video Using Motion Vector-Based Feature Interpolation“
11:30	Invited speaker: Vladlen Koltun. TBD
12:00	Panel discussion with Cordelia Schmid, Daniel Cremers, Bernt Schiele, Vladlen Koltun and the organizers

Workshop Number: 18

Perceptual Image Restoration and Manipulation (PIRM) Workshop and Challenges

Date: Friday 14th, morning

Room: N1070ZG

Organizers: Lihi Zelnik-Manor, Tomer Michaeli, Radu Timofte, Roey Mechrez, Yochai Blau, Andrey Ignatov, Antonio Robles-Kelly, Mehrdad Shoeiby

SCHEDULE

08:30	Opening
08:40 – 09:05	Invited talk – Jiaya Jia (CUHK) "Two Paths to Image Restoration: Deep Learning and Not Using It"
09:05 – 09:30	Invited talk – Eli Shechtman (Adobe Research) "Image Manipulation on the Natural Image Manifold"
09:30 – 10:00	Spotlights
10:00 – 10:45	Poster session and coffee break
10:45 – 11:10	Invited talk – Kavita Bala (Cornell)
11:10 – 11:35	Invited talk – Tomer Michaeli (Technion) "The Perception-Distortion Tradeoff"
11:35 – 12:05	The PIRM Challenges on: (a) Perceptual Super-resolution, (b) Perceptual Enhancement on Smartphones, and (c) Spectral Image Super-resolution
12:05 – 12:30	Invited talk – William Freeman (MIT, Google AI)
12:30	Closing

Workshop Number: 13

1st Large-scale Video Object Segmentation Challenge

Date: Friday 14th, afternoon

Room: 1200 Carl von Linde Hall

Organizers: Ning Xu, Linjie Yang, Yuchen Fan, Jianchao Yang Weiyao Lin, Michael Ying Yang, Brian Price, Jiebo Luo, Thomas Huang

SCHEDULE

12:30 – 13:30	Poster Session (Shared by the workshop "The Third International Workshop on Video Segmentation")
13:30 – 14:00	Welcome and Introduction
14:00 – 14:40	Invited Talk
14:40 – 14:55	4th place team presentation
14:55 – 15:10	Break
15:10 – 15:50	Invited Talk
15:50 – 16:05	3rd place team presentation
16:05 – 16:20	2nd place team presentation
16:20 – 16:35	1st place team presentation
16:35 – 17:00	Closing Remarks and Awards
17:00 – 18:00	Poster Session (Shared by the workshop "The Third International Workshop on Video Segmentation")

Workshop Number: 31

PeopleCap 2018: Capturing and Modelling Human Bodies, Hands and Faces

Date: Friday 14th, afternoon

Room: N1179

Organizers: Gerard Pons-Moll, Jonathan Taylor

SCHEDULE

13:30 – 13:40	Welcome and Introduction
13:40 – 14:20	Invited Talk by Lourdes Agapito
14:20 – 15:00	Invited Talk by Adrian Hilton
15:00 – 16:40	Invited Talk by Franziska Mueller
16:40 – 17:10	Poster Session and Coffee Break
17:10 – 17:40	Invited Talk by Yaser Sheikh
17:40 – 18:10	Invited Talk by Stefanie Wuhrer
18:10	Closing Remarks

Workshop Number: 43

Bias Estimation in Face Analytics (BEFA)

Date: Friday 14th, afternoon

Room: N1070ZG

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

SCHEDULE

13:30	Workshop starts
13:30 – 13:45	Welcome and Competition Winners Announcement
13:45 – 15:30	Invited Talk 1
15:30 – 16:00	Paper Session 1
16:00 – 16:15	Coffee Break
16:15 – 17:00	Invited Talk 2
17:00 – 17:30	Paper Session 2
17:30	Workshop ends

Workshop Number: 11

BioImage Computing

Date: Friday 14th, full day

Room: Theresianum 1601

Organizers: Jens Rittscher, Anna Kreshuk, Florian Jug

SCHEDULE

08:45 – 09:00	Welcome (Jens Rittscher)
09:00 – 09:40	Invited talk (Rene Vidal)
09:40 – 10:20	Mathematical models for bioimage analysis (Virginie Uhlmann)
10:20 – 10:40	A benchmark for epithelial cell tracking (Jan Funke et al)
10:40 – 11:10	Coffee break
11:10 – 11:50	Invited talk (Jean-Christophe Olivo-Marin)
11:50 – 12:10	2D and 3D vascular structures enhancement via multi-scale fractional anisotropy tensor (Haifa F. Alhasson et al)
12:10 – 14:00	Lunch & Poster session
14:00 – 14:40	Invited talk (Rainer Heintzmann)
14:40 – 15:20	Visualising & Interpreting Molecular Landscapes by In-Cell Cryo-Electron Tomography (Julia Mahamid)
15:20 – 15:40	Densely connected stacked U-network for filament segmentation in microscopy images (Yi Liu et al)
15:40 – 16:00	Coffee break
16:00 – 16:20	A fast and scalable pipeline for stain normalization of whole-slide images in histopathology (Milos Stanisavljevic et al)
16:20 – 16:40	Synaptic connectivity detection by (deep) learning signed proximity and pruning (Toufiq Parag et al)
16:40 – 17:00	Multi-level activation for segmentation of hierarchically-nested classes (Marie Piraud et al)

Workshop Number: 15

VizWiz Grand Challenge: Answering Visual Questions from Blind People

Date: Friday 14th, full day

Room: Theresianum 601 EG

Organizers: Danna Gurari, Jeffrey P. Bigham, Kristen Grauman

SCHEDULE

09:00	Opening remarks
09:10 – 09:30	VizWiz: From Visual Question Answering to Supporting Real-World Interactions Jeffrey Bigham
09:30 – 09:50	"Wearable Sensing for Understanding, Forecasting and Assisting Human Activity"; Kris Kitani
09:50 – 10:10	"Forcing Vision and Language Models to Not Just Talk But Also Actually See"; Devi Parikh
10:10 – 10:30	Break
10:30 – 10:50	Overview of challenge, winner announcements, and analysis of results
10:50 – 11:20	Talks by challenge winners
11:20 – 12:30	Poster session
12:30 – 13:45	Lunch
13:45 – 14:05	"Seeing AI: Leveraging Computer Vision to Empower the Blind Community"; Saqib Shaikh
14:05 – 14:25	TBD; Yonatan Wexler
14:25 – 14:45	"Finding and reading scene text without sight" Roberto Manduchi
14:45 – 15:15	Break
15:15 – 15:45	Panel discussion
15:45 – 16:00	Closing discussion & remarks

Workshop Number: 17

3D Reconstruction in the Wild

Date: Friday 14th, full day

Room: N1080ZG

Organizers: Hiroshi Kawasaki, Shohei Nobuhara, Takeshi Masuda, Tomas Pajdla, Akihiro Sugimoto

SCHEDULE

09:00	Workshops starts
09:00 – 9:45	Invited talk 1: Computer Vision, Visual Learning, and 3D Reconstruction, Long Quan
09:45 – 10:30	Invited talk 2: The Long March of 3D Reconstruction: From Tabletop to Outer Space, Andrea Fusiello
10:30 – 11:00	Coffee break
11:00 – 11:45	Invited talk 3: Born in the wild: Self-supervised 3D Face Model Learning, Michael Zollhöfer
11:45 – 12:30	Invited talk 4: Social Perception with Machine Vision, Yaser Shiekh
12:30 – 14:00	Lunch break
14:00 – 15:30	**Poster session:**

1. Deep Depth from Defocus: How Can Defocus Blur Improve 3D Estimation Using Dense Neural Networks? Marcela Carvalho, Bertrand Le Saux, Pauline Trouvé, Andrés Almansa, Frédéric Champagnat

2. Deep Modular Network Architecture for Depth Estimation from Single Indoor Images, Seiya Ito, Naoshi Kaneko, Yuma Shinohara, Kazuhiko Sumi

3. Generative Adversarial Networks for Unsupervise Monocular Depth Prediction, Matteo Poggi, Filippo Aleotti, Fabio Tosi, Stefano Mattoccia

4. Combination of Spatially-Modulated ToF and Structured Light for MPI-Free Depth Estimation, Gianluca Agresti, Pietro Zanuttigh

5. Robust Structured Light System against Subsurface Scattering Effects Achieved by CNN-based Pattern Detection and Decoding Algorithm, Ryo Furukawa, Daisuke Miyazaki, Masashi Baba, Shinsaku Hiura, Hiroshi Kawasaki

6. Robust 3D Pig Measurement in Pig Farm, Kikuhito Kawasue, Kumiko Yoshida

7. SConE: Siamese Constellation Embedding Descriptor for Image Matching, Jacek Komorowski

8. RGB-D SLAM based Incremental Cuboid Modeling, Masashi Mishima, Hideaki Uchiyama, Diego Thomas, Rin-ichiro Taniguchi, Rafael A Roberto, Joao Lima, Veronica Teichrieb

9. Semi-independent Stereo Visual Odometry for Different Field of View Cameras, Trong Phuc Truong, Vincent Nozick, Hideo Saito

10. Improving Thin Structures in Surface Reconstruction from Sparse Point Cloud, Maxime Lhuillier

11. Polygonal Reconstruction of Building Interiors from Cluttered Pointclouds, Inge Coudron, Toon Goedemé, Steven Puttemans

12. Paired 3D Model Generation with Conditional Generative Adversarial Networks, Cihan Öngün, Hilmi Kumdakcı, Alptekin Temizel

13. Predicting Muscle Activity and Joint Angle from Skin Shape, Ryusuke Sagawa, Ko Ayusawa, Yusuke Yoshiyasu, Akihiko Murai

15:30 – 15:45	Coffee break
15:45 – 16:30	Invited talk 5: Depth, Semantics, and Localization for Autonomous Driving Applications, Srikumar Ramalingam
16:30 – 17:15	Invited talk 6: TBA
17:15 – 18:00	Invited talk 7: TBA
18:00	Workshop ends

Workshop Number: 23

6th Workshop on Computer Vision for Road Scene Understanding and Autonomous Driving

Date: Friday 14th, full day

Room: N1189

Organizers: Mathieu Salzmann, José Alvarez, Lars Petersson, Fredrik Kahl, Bart Nabbe

SCHEDULE

08:25	Workshops starts
08:30 – 09:10	Invited talk: Mohan Trivedi (UCSD), "Machine Vision for Human-Centered Autonomous Driving"
09:10 – 09:50	Invited talk: Henning Hamer (Continental), "Continental's eHorizon and Road Database"
09:50 – 11:50	Poster session & Coffee break
11:50 – 12:30	Invited talk: Dariu Gavrila (TU Delft), "EuroCity Persons: A Novel Benchmark for Vulnerable Road User Detection"
12:30 – 14:00	Lunch break
14:00 – 14:40	Invited talk: Arnaud de la Fortelle (Mines ParisTech), "The Perception-Decision Gap"
14:40 – 15:20	Invited talk: Oscar Beijbom (nuTonomy), "The Deep Learning Toolchain for Autonomous Driving"
15:20	Workshop ends

Workshop Number: 24-29

1st Workshop on Interactive and Adaptive Learning in an Open World

Date: Friday 14th, full day

Room: N1090ZG

Organizers: Erik Rodner, Alexander Freytag, Vitto Ferrari, Mario Fritz, Uwe Franke, Terrence Boult, Juergen Gall, Walter Scheirer, Angela Yao

SCHEDULE

09:00	Workshop starts (see https://erodner.github.io/ial2018eccv/ for program updates)
09:00 – 09:10	Workshop opening
09:10 – 09:40	"Incremental learning: a critical view on the current state of affairs"; Tinne Tuytelaars (KU Leuven)
09:45 – 10:15	"Results and Evaluation of the Open-Face Challenge" (http://vast.uccs.edu/Opensetface/); Manuel Günther
10:15 – 10:45	Coffee Break
10:45 – 11:15	"Recognition with unseen compositions and novel environments"; Kristen Grauman (UT Austin)
11:20 – 11:50	"Interactive video segmentation: The DAVIS benchmark and first approaches"; Jordi Pont-Tuset (Google AI)
11:50 – 12:30	Posters (see https://erodner.github.io/ial2018eccv/ for a list of accepted papers)
12:30 – 13:30	Lunch
13:30 – 14:00	Christoph Lampert (IST Austria)
14:05 – 14:35	"Elements of Continuous Learning for Wildlife Monitoring"; Joachim Denzler (Univ. Jena)
14:35 – 14:45	Workshop Closing
14:45	Workshop ends

Workshop Number: 36-68

First Workshop on Computer Vision For Fashion, Art and Design

Date: Friday 14th, full day

Room: Theresianum 602

Organizers: Hui Wu, Negar Rostamzadeh, Leonidas Lefakis, Joy Tang, Rogerio Feris, Tamara Berg, Luba Elliott, Aaron Courville, Chris Pal, Sanja Fidler, Xavier Snelgrove, David Vazquez, Julia Lasserre, Thomas Boquet, Nana Yamazaki

SCHEDULE

9:00 - 9:10	Opening
9:10 -9:50	Invited Talk: Kristen Grauman
9:50 -10:30	Invited Talk: Mario Klingemann
10:30 -11:10	Art Exhibition & Coffee Break
11:10 -11:50	Invited Talk: Tao Mei
11:50-12:30	Invited Talk: Anna Ridler
12:30-13:30	Lunch Break
13:30 -14:10	Invited Talk: Kavita Bala
14:10 -14:50	Spotlight/oral presentations
14:50 -15:50	Poster presentation & Coffee Break
15:50 -16:30	Invited Talk: Aaron Hertzmann
16:30 -17:10	Invited Talk: Larry Davis
17:10 -17:20	Closing Remarks

* Please see the most recent schedule at the workshop website: https://sites.google.com/view/eccvfashion/progr

Workshop Number: 62

TASK-CV Transferring and Adapting Source Knowledge in Computer Vision and VisDA Challenge

Date: Friday 14th, full day

Room: N1095ZG

Organizers: T. Tommasi, D. Vázquez, K. Saenko, B. Usman, X. Peng, J. Hofman, Kaushik, A. M. López, W. Li, F. Orabona

SCHEDULE

08:30 – 08:35	Welcome and agenda presentation
08:35 – 09:10	Invited Talk: Nicolas Courty, University of Bretagne Sud
09:10 – 09:45	Invited Talk: Samory Kpotufe, Princeton University, USA
09:45 – 10:00	Short talk workshop paper
10:00 – 10:15	Short talk workshop paper
10:15 – 11:50	Poster Session and Cofee Break
11:50 – 12:25	Invited Talk: Mingsheng Long, Tsinghua University, China
12:25 – 12:35	Workshop best paper & honorable mention announcement.
12:35 – 14:00	Lunch
14:00 – 14:40	Invited Talk: Ming-Yu Liu, VIDIA Research, USA
14:40 - 14:55	VisDA challenge introduction
14:55 - 15:05	VisDA Open Set Challenge: Honorable Mention Talk
15:05 - 15:15	VisDA Open Set Challenge: Runner-Up Talk
15:15 - 15:25	VisDA Open Set Challenge: Winner Talk
15:25 - 16:00	Cofee Break
16:00 - 16:10	VisDA Detection Challenge: Honorable Mention Talk
16:10 - 16:20	VisDA Detection Challenge: Runner-Up Talk
16:20 - 16:30	VisDA Detection Challenge: Winner Talk
16:30	Best paper announcement and Workshop closing

Workshop Number: 63

The 3rd Workshop on Geometry Meets Deep Learning

Date: Friday 14th, full day

Room: 2750 Karl Marx von Bauernfeind

Organizers: Xiaowei Zhou, Emanuele Rodolà, Jonathan Masci, Kosta Derpanis

SCHEDULE

08:20	Opening
08:30 – 09:00	Invited talk: Sanja Fidler (University of Toronto and NVIDIA AI Research)
09:00 – 09:30	Invited talk: Eric Brachmann (University of Heidelberg)
09:30 – 09:40	Oral presentation: Learning to infer 3D shape using pointillism, Olivia Wiles (University of Oxford), Andrew Zisserman (University of Oxford)
09:40 – 09:50	Oral presentation: High Quality Facial Surface and Texture Synthesis via Generative Adversarial Networks, Ron Slossberg (Technion), Gil Shamai (Technion), Ron Kimmel (Technion)
09:50 – 10:00	Oral presentation: A Simple Approach to Intrinsic Correspondence Learning on Unstructured 3D Meshes, Isaak Lim (RWTH Aachen University), Alexander Dielen (RWTH Aachen University), Marcel Campen (Osnabrück University), Leif Kobbelt (RWTH Aachen University)
10:00 – 11:00	Poster & coffee break
11:00 – 11:30	Invited talk: Leonidas Guibas (Stanford University)
11:30 – 12:00	Invited talk: Kosta Daniilidis (University of Pennsylvania)
12:00 – 13:30	Lunch
13:30 – 14:00	Invited talk: Iasonas Kokkinos (University College London and Facebook AI Research)
14:00 – 14:30	Invited talk: Tomasz Malisiewicz (Magic Leap)

14:30 – 14:40	Oral presentation: Detecting parallel-moving objects in the monocular case employing CNN depth maps, Nolang Fanani (Goethe University Frankfurt), Matthias Ochs (Goethe University Frankfurt), Rudolf Mester (Goethe University Frankfurt)
14:40 – 14:50	Oral presentation: Deep Normal Estimation for Automatic Shading of Hand-Drawn Characters, Matis Hudon (Trinity College Dublin), Mairead Grogan (Trinity College Dublin), Rafael Pagés (Trinity College Dublin), Aljosa Smolic (Trinity College Dublin)
14:50 – 16:00	Poster & coffee break
16:00 – 16:30	Invited talk: Taco Cohen (University of Amsterdam and Qualcomm)
16:30 – 17:00	Invited talk: Michael Bronstein (Università della Svizzera Italiana)
17:00 – 17:30	Invited talk: Noah Snavely (Cornell University and Google)

Posters:

Deep Fundamental Matrix Estimation
Omid Poursaeed (Cornell University), Guandao Yang (Cornell University), Aditya Prakash (Indian Institute of Science), Hanqing Jiang (Cornell University), Quiren Fang (Cornell University), Bharath Hariharan (Cornell University), Serge Belongie (Cornell University)

Evaluation of CNN-based Single-Image Depth Estimation Methods
Tobias Koch (Technical University of Munich), Lukas Liebel (Technical University of Munich), Friedrich Fraundorfer (Graz University of Technology), Marco Körner (Technical University of Munich)

Learning Structure From Motion From Motion
Clément Pinard (ENSTA-ParisTech), Antoine Manzanera (France), David Filliat (ENSTA), Laure Chevalley (Parrot)

Scene Coordinate Regression with Angle-Based Reprojection Loss for Camera Relocalization,
Xiaotian Li (Aalto University), Juha Ylioinas (Aalto University), Jakob Verbeek (INRIA), Juho Kannala (Aalto University, Finland)

Object Pose Estimation from Monocular Image using Multi-View Keypoint

Correspondence
Jogendra Nath Kundu (Indian Institute of Science), Aditya Ganeshan (Indian Institute of Science), Rahul M V (Indian Institute of Science), Venkatesh Babu RADHAKRISHNAN (Indian Institute of Science)

Semi-Supervised Semantic Matching
Zakaria Laskar (Aalto University), Juho Kannala (Aalto University, Finland)

Learning Spectral Transform Network on 3D Surface for Non-rigid Shape Analysis
Ruixuan Yu (Xi'an Jiaotong University), Jian Sun (Xi'an Jiaotong University), Huibin Li (Xian JiaoTong University, China)

Know What Your Neighbors Do: 3D Semantic Segmentation of Point Clouds
Francis Engelmann (RWTH Aachen University, Computer Vision Group), Theodora Kontogianni (RWTH Aachen University, Computer Vision Group), Jonas Schult (RWTH Aachen University)

Multi-Kernel Diffusion CNNs for Graph-Based Learning on Point Clouds
Lasse Hansen (University of Luebeck), Jasper Diesel (Drägerwerk AG & Co. KGaA), Mattias Heinrich (University of Luebeck)

3DContextNet: K-d Tree Guided Hierarchical Learning of Point Clouds Using Local and Global Contextual Cues
Wei Zeng (University of Amsterdam), Theo Gevers (University of Amsterdam)

PosIX-GAN: Generating multiple poses using GAN for Pose-Invariant Face Recognition
Avishek Bhattacharjee (Indian Institute of Technology, Madras), Samik Banerjee (IIT Madras), Sukhendu Das (Indian Institute of Technology, Madras)

Deep Learning for Multi-Path Error Removal in ToF Sensors
Gianluca Agresti (University of Padova), Pietro Zanuttigh (University of Padova)

Attaining human-level performance with atlas location autocontext for anatomical landmark detection in 3D CT data
Alison Q O'Neil (Canon Medical Research Europe)

Workshop Number: 73

Workshop Name: What is Optical Flow for?

Date: Friday 14th, full day

Room: Theresianum 606

Organizers: Laura Sevilla-Lara, Deqing Sun, Jonas Wulff, Fatma Güney

SCHEDULE

9:00	Workshop starts
9:00 – 9:15	Introduction
9:15 – 9:45	Invited Talk: Cordelia Schmid
9:45 – 10:15	Invited Talk: Thomas Brox
10:15 – 10:30	Coffee Break
10:30 – 11:00	Invited Talk: Michael Black
11:00 – 11:30	Invited Talk: Lourdes Agapito
11:30 – 12:00	Invited Talk: Raquel Urtasun
12:00 – 13:00	Lunch Break
13:00 – 13:30	Best Paper presentation
13:30 – 14:00	Invited Talk: Richard Szeliski
14:00 – 14:30	Invited Talk: Kristen Grauman
14:30 – 15:15	Poster Session and Coffee Break
15:15 – 15:45	Invited Talk: Jitendra Malik
15:45 – 16:15	Invited Talk: Bill Freeman
16:15 – 17:00	Round Table Discussion
17:00	Workshop ends

Poster number	Poster title	Authors
P-1A-01	ECO: Efficient Convolutional Network for Online Video Understanding	Mohammadreza Zolfaghari*, University of Freiburg; kamaljeet singh, University of Freiburg; Thomas Brox, University of Freiburg
P-1A-02	Learning to Anonymize Faces for Privacy Preserving Action Detection	Zhongzheng Ren*, University of California, Davis; Yong Jae Lee, University of California, Davis; Michael Ryoo, Indiana University
P-1A-03	Adversarial Open-World Person Re-Identification	Xiang Li, Sun Yat-sen University; Ancong Wu, Sun Yat-sen University; Jason Wei Shi Zheng*, Sun Yat Sen University
P-1A-04	Graph R-CNN for Scene Graph Generation	Jianwei Yang*, Georgia Institute of Technology; Jiasen Lu, Georgia Institute of Technology; Stefan Lee, Georgia Institute of Technology; Dhruv Batra, Georgia Tech & Facebook AI Research; Devi Parikh, Georgia Tech & Facebook AI Research
P-1A-05	Contemplating Visual Emotions: Understanding and Overcoming Dataset Bias	Rameswar Panda*, UC Riverside; Jianming Zhang, Adobe Research; Haoxiang Li, Adobe; Joon-Young Lee, Adobe Research; Xin Lu, Adobe; Amit Roy-Chowdhury , University of California, Riverside, USA
P-1A-06	Graph Adaptive Knowledge Transfer for Unsupervised Domain Adaptation	Zhengming Ding*, Northeastern University; Sheng Li, Adobe Research; Ming Shao, University of Massachusetts Dartmouth; YUN FU, Northeastern University
P-1A-07	Deep Recursive HDRI: Inverse Tone Mapping using Generative Adversarial Networks	Siyeong Lee, Sogang University; Gwon Hwan An, Sogang University; Suk-Ju Kang*, Nil
P-1A-08	Deep Cross-Modal Projection Learning for Image-Text Matching	Ying Zhang*, Dalian University of Technology; Huchuan Lu, Dalian University of Technology

P-1A-09	Composition Loss for Counting, Density Map Estimation and Localization in Dense Crowds	Haroon Idrees*, Carnegie Mellon University; Muhammad Tayyab, UCF; Kishan Athrey, UCF; Mubarak Shah, University of Central Florida; Dong Zhang, University of Central Florida, USA
P-1A-10	Person Search by Multi-Scale Matching	Xu Lan*, Queen Mary University of London; Xiatian Zhu, Queen Mary University, London, UK; Shaogang Gong, Queen Mary University of London
P-1A-11	Efficient 6-DoF Tracking of Handheld Objects from an Egocentric Viewpoint	Rohit Pandey, Google; Pavel Pidlypenskyi, Google; Shuoran Yang, Google; Christine Kaeser-Chen*, Google
P-1A-12	Deep Video Generation, Prediction and Completion of Human Action Sequences	Chunyan Bai, Hong Kong University of Science and Technology; Haoye Cai*, Hong Kong University of Science and Technology; Yu-Wing Tai, Tencent YouTu; Chi-Keung Tang, Hong Kong University of Science and Technology
P-1A-13	Efficient Uncertainty Estimation for Semantic Segmentation in Videos	Po-Yu Huang*, National Tsing Hua University; Wan-Ting Hsu, National Tsing Hua University; Chun-Yueh Chiu, National Tsing Hua University; Tingfan Wu, Umbo Computer Vision; Min Sun, NTHU
P-1A-14	DeepKSPD: Learning Kernel-matrix-based SPD Representation for Fine-grained Image Recognition	Melih Engin, university of wollongong; Lei Wang*, University of Wollongong, Australia; Luping Zhou, University of Wollongong, Australia; Xinwang Liu, National University of Defense Technology
P-1A-15	From Face Recognition to Models of Identity: A Bayesian Approach to Learning about Unknown Identities from Unsupervised Data	Daniel Castro*, Imperial College London; Sebastian Nowozin, Microsoft Research Cambridge

P-1A-16	ShapeStacks: Learning Vision-Based Physical Intuition for Generalised Object Stacking	Oliver Groth*, Oxford Robotics Insitute; Fabian Fuchs, Oxford Robotics Insitute; Andrea Vedaldi, Oxford University; Ingmar Posner, Oxford
P-1A-17	Fast and Precise Camera Covariance Computation for Large 3D Reconstruction	Michal Polic*, Czech Technical University in Prague; Wolfgang Foerstner, University Bonn; Tomas Pajdla, Czech Technical University in Prague
P-1A-18	Inner Space Preserving Generative Pose Machine	Shuangjun Liu, Northeastern University; Sarah Ostadabbas*, Northeastern University
P-1A-19	CTAP: Complementary Temporal Action Proposal Generation	Jiyang Gao*, USC; Kan Chen, University of Southern California, USA; Ram Nevatia, U of Southern California
P-1A-20	Learning to Reenact Faces via Boundary Transfer	Wayne Wu, SenseTime Research; Yunxuan Zhang, sensetime research; Cheng Li*, SenseTime Research; Chen Qian, SenseTime; Chen Change Loy, Chinese University of Hong Kong
P-1A-21	Fast and Accurate Intrinsic Symmetry Detection	Rajendra Nagar*, Indian Institute of Technology Gandhinagar; Shanmuganathan Raman, IIT Gandhinagar
P-1A-22	Fictitious GAN: Training GANs with Historical Models	Yin Xia*, Northwestern University; Xu Chen, Northwestern University; Hao Ge, Northwestern University; Ying Wu, Northwestern University; Randall Berry, Northwestern University
P-1A-23	Audio-Visual Event Localization in Unconstrained Videos	Yapeng Tian*, University of Rochester; Jing Shi, University of Rochester; Bochen Li, University of Rochester; Zhiyao Duan, Unversity of Rochester; Chenliang Xu, University of Rochester
P-1A-24	Tackling 3D ToF Artifacts Through Learning and the FLAT Dataset	Qi Guo, Harvard University; Iuri Frosio*, NVIDIA; Orazio Gallo, NVIDIA Research; Todd Zickler, Harvard University; Kautz Jan, NVIDIA

P-1A-25	Self-Calibrating Isometric Non-Rigid Structure-from-Motion	Shaifali Parashar*, CNRS; Adrien Bartoli, Université Clermont Auvergne; Daniel Pizarro, Universidad de Alcala
P-1A-26	Semi-Supervised Deep Learning with Memory	Yanbei Chen*, Queen Mary University of London; Xiatian Zhu, Queen Mary University, London, UK; Shaogang Gong, Queen Mary University of London
P-1A-27	Question-Guided Hybrid Convolution for Visual Question Answering	Gao Peng*, Chinese university of hong kong; Hongsheng Li, Chinese University of Hong Kong; Shuang Li, The Chinese University of Hong Kong; Pan Lu, Tsinghua University; Yikang LI, The Chinese University of Hong Kong; Steven Hoi, SMU; Xiaogang Wang, Chinese University of Hong Kong, Hong Kong
P-1A-28	Rolling Shutter Pose and Ego-motion Estimation using Shape-from-Template	Yizhen Lao*, Université Clermont Auvergne; Omar Ait-Aider, Université Clermont Auvergne; Adrien Bartoli, Université Clermont Auvergne
P-1A-29	Semi-Dense 3D Reconstruction with a Stereo Event Camera	Yi Zhou*, The Australian National University; Guillermo Gallego, University of Zurich; Henri Rebecq, University of Zurich; Laurent Kneip, ShanghaiTech University; HONGDONG LI, Australian National University, Australia; Davide Scaramuzza, University of Zurich& ETH Zurich, Switzerland
P-1A-30	Local Orthogonal-Group Testing	Ahmet Iscen*, Czech Technical University; Ondrej Chum, Vision Recognition Group, Czech Technical University in Prague
P-1A-31	Temporal Relational Reasoning in Videos	Bolei Zhou*, MIT; Alex Andonian, Massachusetts Institute of Technology; Aude Oliva, MIT; Antonio Torralba, MIT

P-1A-32	Deep High Dynamic Range Imaging with Large Foreground Motions	Shangzhe Wu*, HKUST; Jiarui Xu, Hong Kong University of Science and Technology (HKUST); Yu-Wing Tai, Tencent YouTu; Chi-Keung Tang, Hong Kong University of Science and Technology
P-1A-33	Geometric Constrained Joint Lane Segmentation and Lane Boundary Detection	Jie Zhang*, Shanghai Jiao Tong University; Yi Xu, Shanghai Jiao Tong University; Bingbing Ni, Shanghai Jiao Tong University; Zhenyu Duan, Shanghai Jiao Tong University
P-1A-34	Attributes as Operators	Tushar Nagarajan*, UT Austin; Kristen Grauman, University of Texas
P-1A-35	Textual Explanations for Self-Driving Vehicles	Jinkyu Kim*, UC Berkeley; Anna Rohrbach, UC Berkeley; Trevor Darrell, UC Berkeley; John Canny, UC Berkeley; Zeynep Akata, University of Amsterdam
P-1A-36	Generative Domain-Migration Hashing for Sketch-to-Image Retrieval	Jingyi Zhang*, University of Electronic Science and Technology of China; Fumin Shen, UESTC; Li Liu, the inception institute of artificial intelligence; Fan Zhu, the inception institute of artificial intelligence ; Mengyang Yu, ETH Zurich; Ling Shao, Inception Institute of Artificial Intelligence; Heng Tao Shen, University of Electronic Science and Technology of China (UESTC); Luc Van Gool, ETH Zurich
P-1A-37	Recurrent Fusion Network for Image captioning	Wenhao Jiang*, Tencent AI Lab; Lin Ma, Tencent AI Lab; Yu-Gang Jiang, Fudan University; Wei Liu, Tencent AI Lab; Tong Zhang, Tecent AI Lab
P-1A-38	Attention-based Ensemble for Deep Metric Learning	Wonsik Kim*, Samsung Electronics; Bhavya Goyal, Samsung Electronics; Kunal Chawla, Samsung Electronics; Jungmin Lee, Samsung Electronics; Keunjoo Kwon, Samsung Electronics

P-1A-39	Egocentric Activity Prediction via Event Modulated Attention	Yang Shen*, Shanghai Jiao Tong University; Bingbing Ni, Shanghai Jiao Tong University; Zefan Li, Shanghai Jiao Tong University; Ning Zhuang, Shanghai Jiao Tong University
P-1A-40	A+D Net: Training a Shadow Detector with Adversarial Shadow Attenuation	Hieu Le*, Stony Brook University; Tomas F Yago Vicente, Stony Brook University; Vu Nguyen, Stony Brook University; Minh Hoai Nguyen, Stony Brook University; Dimitris Samaras, Stony Brook University
P-1A-41	Stereo Vision-based Semantic 3D Object and Ego-motion Tracking for Autonomous Driving	Peiliang LI*, HKUST Robotics Institute; Tong QIN, HKUST Robotics Institute; Shaojie Shen, HKUST
P-1A-42	End-to-end View Synthesis for Light Field Imaging with Pseudo 4DCNN	Yunlong Wang*, Center for Research on Intelligent Perception and Computing (CRIPAC) National Laboratory of Pattern Recognition (NLPR) Institute of Automation, Chinese Academy of Sciences (CASIA) ; Fei Liu, Center for Research on Intelligent Perception and Computing (CRIPAC) National Laboratory of Pattern Recognition (NLPR) Institute of Automation, Chinese Academy of Sciences (CASIA); Zilei Wang, University of Science and Technology of China; Guangqi Hou, Center for Research on Intelligent Perception and Computing (CRIPAC) National Laboratory of Pattern Recognition (NLPR) Institute of Automation, Chinese Academy of Sciences (CASIA); Zhenan Sun, Chinese of Academy of Sciences; Tieniu Tan, NLPR, China
P-1A-43	Robust image stitching using multiple registrations	Charles Herrmann, Cornell; Chen Wang, Google Research; Richard Bowen, Cornell; Mike Krainin, Google; Ce Liu, Google; Bill Freeman, MIT; Ramin Zabih*, Cornell Tech/Google Research

P-1A-44	Fast Multi-fiber Network for Video Recognition	Yunpeng Chen*, National University of Singapore; Yannis Kalantidis, Facebook Research, USA; Jianshu Li, NUS; Yan Shuicheng, National University of Singapore; Jiashi Feng, NUS
P-1A-45	TBN: Convolutional Neural Network with Ternary Inputs and Binary Weights	Diwen Wan*, University of Electronic Science and Technology of China; Fumin Shen, UESTC; Li Liu, the inception institute of artificial intelligence; Fan Zhu, the inception institute of artificial intelligence ; Jie Qin, ETH Zurich; Ling Shao, Inception Institute of Artificial Intelligence; Heng Tao Shen, University of Electronic Science and Technology of China (UESTC)
P-1A-46	Contextual Based Image Inpainting: Infer, Match and Translate	Yuhang Song*, USC; Chao Yang, University of Southern California; Zhe Lin, Adobe Research; Xiaofeng Liu, Carnegie Mellon University; Hao Li, Pinscreen/University of Southern California/USC ICT; Qin Huang, University of Southern California; C.-C. Jay Kuo, USC
P-1A-47	Deep Fundamental Matrix Estimation	Rene Ranftl*, Intel Labs; Vladlen Koltun, Intel Labs
P-1A-48	Joint Person Segmentation and Identification in Synchronized First- and Third-person Videos	Mingze Xu*, Indiana University; Chenyou Fan, JD.com; Yuchen Wang, Indiana University; Michael Ryoo, Indiana University; David Crandall, Indiana University
P-1A-49	Linear Span Network for Object Skeleton Detection	Chang Liu*, University of Chinese Academy of Sciences; Wei Ke, University of Chinese Academy of Sciences; Fei Qin, University of Chinese Academy of Sciences; Qixiang Ye, University of Chinese Academy of Sciences, China
P-1A-50	Category-Agnostic Semantic Keypoint Representations in Canonical Object Views	Xingyi Zhou*, The University of Texas at Austin; Arjun Karpur, The University of Texas at Austin; Linjie Luo, Snap Inc; Qixing Huang, The University of Texas at Austin

P-1A-51	Where are the blobs: Counting by Localization with Point Supervision	Issam Hadj Laradji*, University of British Columbia (UBC); Negar Rostamzadeh, Element AI; Pedro Pinheiro, EPFL; David Vazquez, Element AI; Mark Schmidt, University of British Columbia
P-1A-52	A Hybrid Model for Identity Obfuscation by Face Replacement	Qianru Sun*, National University of Singapore; Ayush Tewari, Max Planck Institute for Informatics; Weipeng Xu, MPII; Mario Fritz, Max-Planck-Institut für Informatik; Christian Theobalt, MPI Informatik; Bernt Schiele, MPI
P-1A-53	Exploring the Limits of Supervised Pretraining	Dhruv Mahajan, Facebook; Ross Girshick*, Facebook AI Research (FAIR); Vignesh Ramanathan, Facebook; Kaiming He, Facebook Inc., USA; Manohar Paluri, Facebook; Yixuan Li, Facebook Research; Ashwin Bharambe, Facebook; Laurens van der Maaten, Facebook AI Research
P-1A-54	TrackingNet: A Large-Scale Dataset and Benchmark for Object Tracking in the Wild	Matthias Müller*, King Abdullah University of Science and Technology (KAUST); Adel Bibi, KAUST; Silvio Giancola, KAUST; Salman Al-Subaihi, KAUST; Bernard Ghanem, KAUST
P-1A-55	Unpaired Image Captioning by Language Pivoting	Jiuxiang Gu*, Nanyang Technological University; Shafiq Joty, Nanyang Technological University; Jianfei Cai, Nanyang Technological University; Gang Wang, Alibaba Group
P-1A-56	Pairwise Relational Networks for Face Recognition	Bong-Nam Kang*, POSTECH
P-1A-57	DeepPhys: Video-Based Physiological Measurement Using Convolutional Attention Networks	Weixuan Chen*, MIT Media Lab; Daniel McDuff, Microsoft Research

P-1A-58	Semantic Match Consistency for Long-Term Visual Localization	Carl Toft*, Chalmers; Erik Stenborg, Chalmers University; Lars Hammarstrand, Chalmers university of technology; Lucas Brynte, Chalmers University of Technology; Marc Pollefeys, ETH Zurich; Torsten Sattler, ETH Zurich; Fredrik Kahl, Chalmers
P-1A-59	Grounding Visual Explanations	Lisa Anne Hendricks*, Uc berkeley; Ronghang Hu, University of California, Berkeley; Trevor Darrell, UC Berkeley; Zeynep Akata, University of Amsterdam
P-1A-60	Cross-Modal Hamming Hashing	Yue Cao, Tsinghua University; Mingsheng Long*, Tsinghua University; Bin Liu, Tsinghua University; Jianmin Wang, Tsinghua University, China
P-1A-61	A Modulation Module for Multi-task Learning with Applications in Image Retrieval	Xiangyun Zhao*, Northwestern University; Haoxiang Li, Adobe; Xiaohui Shen, Adobe Research; Xiaodan Liang, Carnegie Mellon University; Ying Wu, Northwestern University
P-1A-62	Open-World Stereo Video Matching with Deep RNN	Yiran Zhong*, Australian National University; HONGDONG LI, Australian National University, Australia; Yuchao Dai, Northwestern Polytechnical University
P-1A-63	Deblurring Natural Image Using Super-Gaussian Fields	Yuhang Liu, Wuhan University; Wenyong Dong*, Wuhan University; Dong Gong, Northwestern Polytechnical University & The University of Adelaide; Lei Zhang, The unversity of Adelaide; Qinfeng Shi, University of Adelaide
P-1A-64	Diverse and Coherent Paragraph Generation from Images	Moitreya Chatterjee*, University of Illinois at Urbana Champaign; Alexander Schwing, UIUC
P-1A-65	Learning Compression from limited unlabeled Data	Xiangyu He*, Chinese Academy of Sciences; Jian Cheng, Chinese Academy of Sciences, China

P-1A-66	Deep Video Quality Assessor: From Spatio-temporal Visual Sensitivity to A Convolutional Neural Aggregation Network	Woojae Kim*, Yonsei University; Jongyoo Kim, Yonsei University; Sewoong Ahn, Yonsei University; Jinwoo Kim, Yonsei University; Sanghoon Lee, Yonsei University, Korea
P-1A-67	Product Quantization Network for Fast Image Retrieval	Tan Yu*, Nanyang Technological University; Junsong Yuan, State University of New York at Buffalo, USA; CHEN FANG, Adobe Research, San Jose, CA; Hailin Jin, Adobe Research
P-1A-68	Factorizable Net: An Efficient Subgraph-based Framework for Scene Graph Generation	Yikang LI*, The Chinese University of Hong Kong; Bolei Zhou, MIT; Yawen Cui, National University of Defense Technology ; Jianping Shi, Sensetime Group Limited; Xiaogang Wang, Chinese University of Hong Kong, Hong Kong; Wanli Ouyang, CUHK
P-1A-69	C-WSL: Count-guided Weakly Supervised Localization	Mingfei Gao*, University of Maryland; Ang Li, Google DeepMind; Ruichi Yu, University of Maryland, College Park; Vlad Morariu, Adobe Research; Larry Davis, University of Maryland
P-1A-70	The Sound of Pixels	Hang Zhao*, Massachusetts Institute of Technology; Chuang Gan, MIT; Andrew Rouditchenko, MIT; Carl Vondrick, MIT; Josh McDermott, Massachusetts Institute of Technology; Antonio Torralba, MIT
P-1A-71	Unsupervised Video Object Segmentation using Motion Saliency-Guided Spatio-Temporal Propagation	Yuan-Ting Hu*, University of Illinois at Urbana-Champaign; Jia-Bin Huang, Virginia Tech; Alexander Schwing, UIUC
P-1A-72	Good Line Cutting: towards Accurate Pose Tracking of Line-assisted VO/VSLAM	Yipu Zhao*, Georgia Institute of Technology; Patricio Vela, Georgia Institute of Technology
P-1A-73	Bi-box Regression for Pedestrian Detection and Occlusion Estimation	CHUNLUAN ZHOU*, Nanyang Technological University; Junsong Yuan, State University of New York at Buffalo, USA

P-1A-74	Unveiling the Power of Deep Tracking	Goutam Bhat*, Linkoping University; Joakim Johnander, Linköping University; Martin Danelljan, Linkoping University; Fahad Shahbaz Khan, Linköping University; Michael Felsberg, Linköping University
P-1A-75	Multi-Scale Structure-Aware Network for Human Pose Estimation	Lipeng Ke*, University of Chinese Academy of Sciences; Ming-Ching Chang, Albany University; Honggang Qi, University of Chinese Academy of Sciences; Siwei Lyu, University at Albany
P-1A-76	Neural Graph Matching Networks for Fewshot 3D Action Recognition	Michelle Guo*, Stanford University; Edward Chou, Stanford University; De-An Huang, Stanford University; Shuran Song, Princeton; Serena Yeung, Stanford University; Li Fei-Fei, Stanford University
P-1A-77	Objects that Sound	Relja Arandjelovi?*, DeepMind; Andrew Zisserman, University of Oxford
P-1A-78	Discriminative Region Proposal Adversarial Networks for High-Quality Image-to-Image Translation	Chao Wang, Ocean University of China; Haiyong Zheng*, Ocean University of China; Zhibin Yu, Ocean University of China; Ziqiang Zheng, Ocean University of China; Zhaorui Gu, Ocean University of China; Bing Zheng, Ocean University of China
P-1A-79	SaaS: Speed as a Supervisor for Semi-supervised Learning	Safa Cicek*, UCLA; Alhussein Fawzi, UCLA; Stefano Soatto, UCLA
P-1A-80	Adaptive Affinity Field for Semantic Segmentation	Tsung-Wei Ke, UC Berkeley / ICSI; Jyh-Jing Hwang*, UC Berkeley / ICSI; Ziwei Liu, UC Berkeley / ICSI; Stella Yu, UC Berkeley / ICSI
P-1A-81	Semi-convolutional Operators for Instance Segmentation	Samuel Albanie*, University of Oxford; Andrea Vedaldi, Oxford University; David Novotny, Oxford University; Diane Larlus, Naver Labs Europe

P-1A-82	Effective Use of Synthetic Data for Urban Scene Semantic Segmentation	Fatemeh Sadat Saleh*, Australian National University (ANU); Mohammad Sadegh Aliakbarian, Data61; Mathieu Salzmann, EPFL; Lars Petersson, Data61/CSIRO; Jose Manuel Alvarez, Toyota Research Institute
P-1A-83	Shape correspondences from learnt template-based parametrization	Thibault Groueix*, École des ponts ParisTech; Bryan Russell, Adobe Research; Mathew Fisher, Adobe Research; Vladimir Kim, Adobe Research; Mathieu Aubry, École des ponts ParisTech
P-1A-84	TextSnake: A Flexible Representation for Detecting Text of Arbitrary Shapes	Shangbang Long, Peking University; Jiaqiang Ruan, Peking University; Wenjie Zhang, Peking University; Xin He*, Megvii; Wenhao Wu, Megvii; Cong Yao, Megvii
P-1A-85	How good is my GAN?	Konstantin Shmelkov*, Inria; Cordelia Schmid, INRIA; Karteek Alahari, Inria
P-1A-86	Deep Generative Models for Weakly-Supervised Multi-Label Classification	Hong-Min Chu*, National Taiwan University; Chih-Kuan Yeh, Carnegie Mellon University; Yu-Chiang Frank Wang, National Taiwan University
P-1A-87	Attention-GAN for Object Transfiguration in Wild Images	Xinyuan Chen*, Shanghai Jiao Tong University; Chang Xu, University of Sydney; Xiaokang Yang, Shanghai Jiao Tong University of China; Dacheng Tao, University of Sydney
P-1A-88	Skeleton-Based Action Recognition with Spatial Reasoning and Temporal Stack Learning	Chenyang Si*, Institute of Automation, Chinese Academy of Sciences; Ya Jing, Institute of Automation, Chinese Academy of Sciences; wei wang, Institute of Automation Chinese Academy of Sciences; Liang Wang, NLPR, China; Tieniu Tan, NLPR, China
P-1A-89	Diverse Image-to-Image Translation via Disentangled Representations	Hsin-Ying Lee*, University of California, Merced; Hung-Yu Tseng, University of California, Merced; Maneesh Singh, Verisk Analytics; Jia-Bin Huang, Virginia Tech; Ming-Hsuan Yang, University of California at Merced

P-1A-90	Convolutional Networks with Adaptive Computation Graphs	Andreas Veit*, Cornell University; Serge Belongie, Cornell University
P-1B-01	Learning to Separate Object Sounds by Watching Unlabeled Video	Ruohan Gao*, University of Texas at Austin; Rogerio Feris, IBM Research; Kristen Grauman, University of Texas
P-1B-02	Learning-based Video Motion Magni_x000c_cation	Tae-Hyun Oh, MIT CSAIL; Ronnachai Jaroensri*, MIT CSAIL; Changil Kim, MIT CSAIL; Mohamed A. Elghareb, Qatar Computing Research Institute; Fredo Durand, MIT; Bill Freeman, MIT; Wojciech Matusik, Adobe
P-1B-03	Light Structure from Pin Motion: Simple and Accurate Point Light Calibration for Physics-based Modeling	Hiroaki Santo*, Osaka University; Michael Waechter, Osaka University; Masaki Samejima, Osaka University; Yusuke Sugano, Osaka University; Yasuyuki Matsushita, Osaka University
P-1B-04	Video Object Segmentation with Joint Re-identification and Attention-Aware Mask Propagation	Xiaoxiao Li*, The Chinese University of Hong Kong; Chen Change Loy, Chinese University of Hong Kong
P-1B-05	Coded Two-Bucket Cameras for Computer Vision	Mian Wei, University of Toronto; Navid Navid Sarhangnejad, University of Toronto; Zhengfan Xia, University of Toronto; Nikola Katic, University of Toronto; Roman Genov, University of Toronto; Kyros Kutulakos*, University of Toronto
P-1B-06	Multimodal Unsupervised Image-to-image Translation	Xun Huang*, Cornell University; Ming-Yu Liu, NVIDIA; Serge Belongie, Cornell University; Kautz Jan, NVIDIA

P-1B-07	Learning to Detect and Track Visible and Occluded Body Joints in a Virtual World	Matteo Fabbri, University of Modena and Reggio Emilia; Fabio Lanzi*, University of Modena and Reggio Emilia; SIMONE CALDERARA, University of Modena and Reggio Emilia, Italy; Andrea Palazzi, University of Modena and Reggio Emilia; ROBERTO VEZZANI, University of Modena and Reggio Emilia, Italy; Rita Cucchiara, Universita Di Modena E Reggio Emilia
P-1B-08	Local Spectral Graph Convolution for Point Set Feature Learning	Chu Wang*, McGill University; Babak Samari, McGill University; Kaleem Siddiqi, McGill University
P-1B-09	Meta-Tracker: Fast and Robust Online Adaptation for Visual Object Trackers	Eunbyung Park*, UNC-CHAPEL HILL; Alex Berg, University of North Carolina, USA
P-1B-10	VSO: Visual Semantic Odometry	Konstantinos-Nektarios Lianos, Geomagical Labs, Inc; Johannes Schoenberger, ETH Zurich; Marc Pollefeys, ETH Zurich; Torsten Sattler*, ETH Zurich
P-1B-11	Progressive Lifelong Learning by Distillation and Retrospection	Saihui Hou*, University of Science and Technology of China; Xinyu Pan, MMLAB, CUHK; Chen Change Loy, Chinese University of Hong Kong; Dahua Lin, The Chinese University of Hong Kong
P-1B-12	Spatio-Temporal Channel Correlation Networks for Action Classification	Ali Diba*, KU Leuven; Mohsen Fayyaz, University of Bonn; Vivek Sharma, Karlsruhe Institute of Technology; Mohammad Arzani, Sensifai; Rahman Yousefzadeh, sensifai; Jürgen Gall, University of Bonn; Luc Van Gool, ETH Zurich
P-1B-13	Long-term Tracking in the Wild: a Benchmark	Efstratios Gavves, University of Amsterdam ; Luca Bertinetto*, University of Oxford; Joao Henriques, University of Oxford; Andrea Vedaldi, Oxford University; Philip Torr, University of Oxford; Ran Tao, University of Amsterdam; Jack Valmadre, Oxford

P-1B-14	Online Detection of Action Start in Untrimmed, Streaming Videos	Zheng Shou*, Columbia University; Junting Pan, Columbia University ; Jonathan Chan, Columbia University; Kazuyuki Miyazawa, Mitsubishi Electric; Hassan Mansour, Mitsubishi Electric Research Laboratories (MERL); Anthony Vetro, Mitsubishi Electric Research Lab; Xavier Giro-i-Nieto, Universitat Politecnica de Catalunya; Shih-Fu Chang, Columbia University
P-1B-15	Two Stream Pose Transfer Guided by Dense Pose Estimation	Natalia Neverova*, Facebook AI Research; Alp Guler, INRIA; Iasonas Kokkinos, Facebook, France
P-1B-16	Simultaneous 3D Reconstruction for Water Surface and Underwater Scene	Yiming Qian*, University of Alberta; Yinqiang Zheng, National Institute of Informatics; Minglun Gong, Memorial University; Herb Yang, University of Alberta
P-1B-17	Stereo gaze: Inferring novel 3D locations from gaze in monocular video	Ernesto Brau, CiBO Technologies; Jinyan Guan, UC San Diego; Tanya Jeffries, U. Arizona; Kobus Barnard*, University of Arizona
P-1B-18	Multi-Scale Context Intertwining for Semantic Segmentation	Di Lin*, Shenzhen University; Yuanfeng Ji, Shenzhen University; Dani Lischinski, The Hebrew University of Jerusalem; Danny Cohen-Or, Tel Aviv University; Hui Huang, Shenzhen University
P-1B-19	Object-centered image stitching	Charles Herrmann, Cornell; Chen Wang, Google Research; Richard Bowen, Cornell; Ramin Zabih*, Cornell Tech/Google Research
P-1B-20	Grassmann Pooling for Fine-Grained Visual Classification	Xing Wei*, Xi'an Jiaotong University; Yihong Gong, Xi'an Jiaotong University; Yue Zhang, Xi'an Jiaotong University; Nanning Zheng, Xi'an Jiaotong University; Jiawei Zhang, City University of Hong Kong
P-1B-21	Diagnosing Error in Temporal Action Detectors	Humam Alwassel*, KAUST; Fabian Caba, KAUST; Victor Escorcia, KAUST; Bernard Ghanem, KAUST

P-1B-22	CGIntrinsics: Better Intrinsic Image Decomposition through Physically-Based Rendering	Zhengqi Li*, Cornell University; Noah Snavely, -
P-1B-23	A Closed-form Solution to Photorealistic Image Stylization	Yijun Li*, University of California, Merced; Ming-Yu Liu, NVIDIA; Xueting Li, University of California, Merced; Ming-Hsuan Yang, University of California at Merced; Kautz Jan, NVIDIA
P-1B-24	Two at Once: Enhancing Learning and Generalization Capacities via IBN-Net	Xingang Pan*, The Chinese University of Hong Kong; Ping Luo, The Chinese University of Hong Kong; Jianping Shi, Sensetime Group Limited; Xiaoou Tang, The Chinese University of Hong Kong
P-1B-25	Collaborative Deep Reinforcement Learning for Multi-Object Tracking	Liangliang Ren, Tsinghua University; Zifeng Wang, Tsinghua University; Jiwen Lu*, Tsinghua University; Qi Tian , The University of Texas at San Antonio; Jie Zhou, Tsinghua University, China
P-1B-26	Single Image Highlight Removal with a Sparse and Low-Rank Reflection Model	Jie Guo*, Nanjing University; Zuojian Zhou, Nanjing University Of Chinese Medicine; Limin Wang, Nanjing University
P-1B-27	Hierarchical Relational Networks for Group Activity Recognition and Retrieval	Mostafa Ibrahim*, Simon Fraser University; Greg Mori, Simon Fraser University
P-1B-28	Towards Human-Level License Plate Recognition	Jiafan Zhuang, University of Science and Technology of China; Zilei Wang*, University of Science and Technology of China
P-1B-29	Stacked Cross Attention for Image-Text Matching	Kuang-Huei Lee*, Microsoft AI and Research; Xi Chen, Microsoft AI and Research; Gang Hua, Microsoft Cloud and AI; Houdong Hu, Microsoft AI and Research; Xiaodong He, JD AI Research

P-1B-30	Deep Discriminative Model for Video Classification	Mohammad Tavakolian*, University of Oulu; Abdenour Hadid, Finland
P-1B-31	The Mutex Watershed: Efficient, Parameter-Free Image Partitioning	Steffen Wolf*, Univertity of Heidelberg; Constantin Pape, University of Heidelberg; Nasim Rahaman, University of Heidelberg; Anna Kreshuk, University of Heidelberg; Ullrich Köthe, University of Heidelberg; Fred Hamprecht, Heidelberg Collaboratory for Image Processing
P-1B-32	Monocular Depth Estimation with Affinity, Vertical Pooling, and Label Enhancement	YuKang Gan*, SUN YAT-SEN University; Xiangyu Xu, Tsinghua University; Wenxiu Sun, SenseTime Research; Liang Lin, SenseTime
P-1B-33	Improved Structure from Motion Using Fiducial Marker Matching	Joseph DeGol*, UIUC; Timothy Bretl, University of Illinois at Urbana-Champaign; Derek Hoiem, University of Illinois at Urbana-Champaign
P-1B-34	Temporal Modular Networks for Retrieving Complex Compositional Activities in Video	Bingbin Liu*, Stanford University; Serena Yeung, Stanford University; Edward Chou, Stanford University; De-An Huang, Stanford University; Li Fei-Fei, Stanford University; Juan Carlos Niebles, Stanford University
P-1B-35	Quantized Densely Connected U-Nets for Efficient Landmark Localization	Zhiqiang Tang*, Rutgers; Xi Peng, Rutgers University; Shijie Geng, Rutgers; Shaoting Zhang, University of North Carolina at Charlotte; Lingfei Wu, IBM T. J. Watson Research Center; Dimitris Metaxas, Rutgers
P-1B-36	Real-to-Virtual Domain Uni_x000c_cation for End-to-End Autonomous Driving	Luona Yang*, Carnegie Mellon University; Xiaodan Liang, Carnegie Mellon University; Eric Xing, Petuum Inc.
P-1B-37	Beyond Part Models: Person Retrieval with Refined Part Pooling (and A Strong Convolutional Baseline)	Yifan Sun*, Tsinghua University; Liang Zheng, Singapore University of Technology and Design; Yi Yang, University of Technology, Sydney; Qi Tian , The University of Texas at San Antonio; Shengjin Wang, Tsinghua University

P-1B-38	Fully-Convolutional Point Networks for Large-Scale Point Clouds	Dario Rethage*, Technical University of Munich, Germany; Johanna Wald, Technical University of Munich; Nassir Navab, TU Munich, Germany; Federico Tombari, Technical University of Munich, Germany
P-1B-39	Real-Time Hair Rendering using Sequential Adversarial Networks	Lingyu Wei*, University of Southern California; Liwen Hu, University of Southern California; Vladimir Kim, Adobe Research; Ersin Yumer, Argo AI; Hao Li, Pinscreen/University of Southern California/USC ICT
P-1B-40	Visual Tracking via Spatially Aligned Correlation Filters Network	mengdan zhang*, Institute of Automation, Chinese Academy of Sciences; qiang wang, Institute of Automation, Chinese Academy of Sciences; Junliang Xing, National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences; Jin Gao, Institute of Automation, Chinese Academy of Sciences; peixi peng, Institute of Automation, Chinese Academy of Sciences; Weiming Hu, Institute of Automation,Chinese Academy of Sciences; Steve Maybank, University of London
P-1B-41	Spatio-temporal Transformer Network for Video Restoration	Tae Hyun Kim*, Max Planck Institute for Intelligent Systems; Mehdi S. M. Sajjadi, Max Planck Institute for Intelligent Systems; Michael Hirsch, Max Planck Institut for Intelligent Systems ; Bernhard Schölkopf, Max Planck Institute for Intelligent Systems
P-1B-42	Value-aware Quantization for Training and Inference of Neural Networks	Eunhyeok Park, Seoul National University; Sungjoo Yoo*, Seoul National University; Peter Vajda, Facebook
P-1B-43	Lambda Twist: An Accurate Fast Robust Perspective Three Point (P3P) Solver	Mikael Persson*, Linköping University; Klas Nordberg, Linköping University

P-1B-44	Programmable Light Curtains	Jian Wang*, Carnegie Mellon University; Joe Bartels, Carnegie Mellon University; William Whittaker, Carnegie Mellon University; Aswin Sankaranarayanan, Carnegie Mellon University; Srinivasa Narasimhan, Carnegie Mellon University
P-1B-45	Monocular Depth Estimation Using Whole Strip Masking and Reliability-Based Refinement	Minhyeok Heo*, Korea University; Jaehan Lee, Korea University; Kyung-Rae Kim, Korea University; Han-Ul Kim, Korea University; Chang-Su Kim, Korea university
P-1B-46	Task-Aware Image Downscaling	Heewon Kim, Seoul National University; Myungsub Choi, Seoul National University; Bee Lim, Seoul National University; Kyoung Mu Lee*, Seoul National University
P-1B-47	Single Image Scene Refocusing using Conditional Adversarial Networks	Parikshit Sakurikar*, IIIT-Hyderabad; Ishit Mehta, IIIT Hyderabad; Vineeth N Balasubramanian, IIT Hyderabad; P. J. Narayanan, IIIT-Hyderabad
P-1B-48	Model-free Consensus Maximization for Non-Rigid Shapes	Thomas Probst*, ETH Zurich; Ajad Chhatkuli , ETHZ; Danda Pani Paudel, ETH Zürich; Luc Van Gool, ETH Zurich
P-1B-49	BSN: Boundary Sensitive Network for Temporal Action Proposal Generation	Tianwei Lin, Shanghai Jiao Tong University; Xu Zhao*, Shanghai Jiao Tong University; Haisheng Su, Shanghai Jiao Tong University; Chongjing Wang, China Academy of Information and Communications Technology; Ming Yang, Shanghai Jiao Tong University
P-1B-50	Materials for Masses: SVBRDF Acquisition with a Single Mobile Phone Image	Zhengqin Li*, UC San Diego; Manmohan Chandraker, UC San Diego; Sunkavalli Kalyan, Adobe Research
P-1B-51	Attentive Semantic Alignment using Offset-Aware Correlation Kernels	Paul Hongsuck Seo*, POSTECH; Jongmin Lee, POSTECH; Deunsol Jung, POSTECH; Bohyung Han, Seoul National University; Minsu Cho, POSTECH

P-1B-52	Deeply Learned Compositional Models for Human Pose Estimation	Wei Tang*, Northwestern University; Pei Yu, Northwestern University; Ying Wu, Northwestern University
P-1B-53	Real-Time Tracking with Discriminative Multi-Domain Convolutional Neural Networks	Ilchae Jung*, POSTECH; Jeany Son, POSTECH; Mooyeol Baek, POSTECH; Bohyung Han, Seoul National University
P-1B-54	Women also Snowboard: Overcoming Bias in Captioning Models	Lisa Anne Hendricks*, Uc berkeley; Anna Rohrbach, UC Berkeley; Kaylee Burns, UC Berkeley; Trevor Darrell, UC Berkeley
P-1B-55	Progressive Structure from Motion	Alex Locher*, ETH Zürich; Michal Havlena, Vuforia, PTC, Vienna; Luc Van Gool, ETH Zurich
P-1B-56	Occlusion-aware R-CNN: Detecting Pedestrians in a Crowd	Shifeng Zhang*, CBSR, NLPR, CASIA; Longyin Wen, GE Global Research; Xiao Bian, GE Global Research; Zhen Lei, NLPR, CASIA, China; Stan Li, National Lab. of Pattern Recognition, China
P-1B-57	Affinity Derivation and Graph Merge for Instance Segmentation	Yiding Liu*, University of Science and Technology of China; Siyu Yang, Beihang University; Bin Li, Microsoft Research Asia; Wengang Zhou, University of Science and Technology of China; Ji-Zeng Xu, Microsoft Research Asia; Houqiang Li, University of Science and Technology of China; Yan Lu, Microsoft Research Asia
P-1B-58	Second-order Democratic Aggregation	Tsung-Yu Lin*, University of Massachusetts Amherst; Subhransu Maji, University of Massachusetts, Amherst; Piotr Koniusz, Data61/CSIRO, ANU
P-1B-59	Improving Sequential Determinantal Point Processes for Supervised Video Summarization	Aidean Sharghi*, University of Central Florida; Boqing Gong, Tencent AI Lab; Ali Borji, University of Central Florida; Chengtao Li, MIT; Tianbao Yang, University of Iowa

P-1B-60	Seeing Deeply and Bi-directionally: A Deep Learning Approach for Single Image Reflection Removal	Jie Yang*, University of Adelaide; Dong Gong, Northwestern Polytechnical University & The University of Adelaide; Lingqiao Liu, University of Adelaide; Qinfeng Shi, University of Adelaide
P-1B-61	Specular-to-Diffuse Translation for Multi-View Reconstruction	Shihao Wu*, University of Bern; Hui Huang, Shenzhen University; Tiziano Portenier, University of Bern; Matan Sela, Technion - Israel Institute of Technology; Danny Cohen-Or, Tel Aviv University; Ron Kimmel, Technion; Matthias Zwicker, University of Maryland
P-1B-62	SEAL: A Framework Towards Simultaneous Edge Alignment and Learning	Zhiding Yu*, NVIDIA; Weiyang Liu, Georgia Tech; Yang Zou, Carnegie Mellon University; Chen Feng, Mitsubishi Electric Research Laboratories (MERL); Srikumar Ramalingam, University of Utah; B. V. K. Vijaya Kumar, CMU, USA; Kautz Jan, NVIDIA
P-1B-63	Question Type Guided Attention in Visual Question Answering	Yang Shi*, University of California, Irvine; Tommaso Furlanello, University of Southern California; Sheng Zha, Amazon Web Services; Anima Anandkumar, Amazon
P-1B-64	Neural Procedural Reconstruction for Residential Buildings	Huayi Zeng*, Washington University in St.Louis; Jiaye Wu, Washington University in St.Louis; Yasutaka Furukawa, Simon Fraser University
P-1B-65	Self-Calibration of Cameras with Euclidean Image Plane in Case of Two Views and Known Relative Rotation Angle	Evgeniy Martyushev*, South Ural State University
P-1B-66	Towards Optimal Deep Hashing via Policy Gradient	Xin Yuan, Tsinghua University; Liangliang Ren, Tsinghua University; Jiwen Lu*, Tsinghua University; Jie Zhou, Tsinghua University, China

P-1B-67	Piggyback: Adapting a Single Network to Multiple Tasks by Learning to Mask Weights	Arun Mallya*, UIUC; Svetlana Lazebnik, UIUC; Dillon Davis, UIUC
P-1B-68	Generating 3D Faces using Convolutional Mesh Autoencoders	Anurag Ranjan*, MPI for Intelligent Systems; Timo Bolkart, Max Planck for Intelligent Systems; Soubhik Sanyal, Max Planck Institute for Intelligent Systems; Michael Black, Max Planck Institute for Intelligent Systems
P-1B-69	ICNet for Real-Time Semantic Segmentation on High-Resolution Images	Hengshuang Zhao, The Chinese University of Hong Kong; Xiaojuan Qi, CUHK; Xiaoyong Shen*, CUHK; Jianping Shi, Sensetime Group Limited; Jia Jiaya, Chinese University of Hong Kong
P-1B-70	Memory Aware Synapses: Learning what (not) to forget	Rahaf Aljundi*, KU Leuven; Francesca babiloni, KU Leuven; Mohamed Elhoseiny, Facebook; Marcus Rohrbach, Facebook AI Research; Tinne Tuytelaars, K.U. Leuven
P-1B-71	Deep Texture and Structure Aware Filtering Network for Image Smoothing	Kaiyue Lu*, Australian National University & Data61-CSIRO; Shaodi You, Data61-CSIRO, Australia; Nick Barnes, CSIRO(Data61)
P-1B-72	Linear RGB-D SLAM for Planar Environments	Pyojin Kim*, Seoul National University; Brian Coltin, NASA Ames Research Center; Hyoun Jin Kim, Seoul National University
P-1B-73	DeepJDOT: Deep Joint distribution optimal transport for unsupervised domain adaptation	Bharath Bhushan Damodaran*, IRISA,Universite de Bretagne-Sud; Benjamin Kellenberger, Wageningen University and Research; Rémi Flamary, Université Côte d Azur; Devis Tuia, Wageningen University and Research; Nicolas Courty, IRISA, Universite Bretagne-Sud

P-1B-74	W-TALC: Weakly-supervised Temporal Activity Localization and Classification	Sujoy Paul*, University of California-Riverside; Sourya Roy, University of California, Riverside; Amit Roy-Chowdhury , University of California, Riverside, USA
P-1B-75	Unsupervised Video Object Segmentation with Motion-based Bilateral Networks	Siyang Li*, University of Southern California; Bryan Seybold, Google Inc.; Alexey Vorobyov, Google Inc.; Xuejing Lei, University of Southern California ; C.-C. Jay Kuo, USC
P-1B-76	Disentangling Factors of Variation with Cycle-Consistent Variational Auto-Encoders	Ananya Harsh Jha*, Indraprastha Institute of Information Technology Delhi; Saket Anand, Indraprastha Institute of Information Technology Delhi; Maneesh Singh, Verisk Analytics; VSR Veeravasarapu, Verisk Analytics
P-1B-77	Mancs: A Multi-task Attentional Network with Curriculum Sampling for Person Re-identification	Cheng Wang, Huazhong Univ. of Science and Technology; Qian Zhang, Horizon Robotics; Chang Huang, Horizon Robotics, Inc.; Wenyu Liu, Huazhong University of Science and Technology; Xinggang Wang*, Huazhong Univ. of Science and Technology
P-1B-78	Multi-view to Novel view: Synthesizing Views via Self-Learned Confidence	Shao-Hua Sun*, University of Southern California; Jacob Huh, Carnegie Mellon University; Yuan-Hong Liao, National Tsing Hua University; Ning Zhang, SnapChat; Joseph Lim, USC
P-1B-79	Part-Activated Deep Reinforcement Learning for Action Prediction	Lei Chen, Tianjin University; Jiwen Lu*, Tsinghua University; Zhanjie Song, Tianjin University; Jie Zhou, Tsinghua University, China
P-1B-80	Online Dictionary Learning for Approximate Archetypal Analysis	Jieru Mei, Microsoft Research Asia; Chunyu Wang*, Microsoft Research asia; Wenjun Zeng, Microsoft Research
P-1B-81	Estimating Depth from RGB and Sparse Sensing	Zhao Chen*, Magic Leap, Inc.; Vijay Badrinarayanan, Magic Leap, Inc.; Gilad Drozdov, Magic Leap, Inc.; Andrew Rabinovich, Magic Leap, Inc.

P-1B-82	Unsupervised Domain Adaptation for Semantic Segmentation via Class-Balanced Self-Training	Yang Zou*, Carnegie Mellon University; Zhiding Yu, NVIDIA; B. V. K. Vijaya Kumar, CMU, USA; Jinsong Wang, General Motors
P-1B-83	Zoom-Net: Mining Deep Feature Interactions for Visual Relationship Recognition	Guojun Yin, University of Science and Technology of China; Lu Sheng, The Chinese University of Hong Kong; Bin Liu, University of Science and Technology of China; Nenghai Yu, University of Science and Technology of China; Xiaogang Wang, Chinese University of Hong Kong, Hong Kong; Chen Change Loy, Chinese University of Hong Kong; Jing Shao*, The Chinese University of Hong Kong
P-1B-84	Joint Camera Spectral Sensitivity Selection and Hyperspectral Image Recovery	Ying Fu*, Beijing Institute of Technology; Tao Zhang, Beijing Institute of Technology; Yinqiang Zheng, National Institute of Informatics; debing zhang, DeepGlint; Hua Huang, Beijing Institute of Technology
P-1B-85	Compositing-aware Image Search	Hengshuang Zhao*, The Chinese University of Hong Kong; Xiaohui Shen, Adobe Research; Zhe Lin, Adobe Research; Sunkavalli Kalyan, Adobe Research; Brian Price, Adobe; Jia Jiaya, Chinese University of Hong Kong
P-1B-86	Zero-shot keyword search for visual speech recognition in-the-wild	Themos Stafylakis*, University of Nottingham; Georgios Tzimiropoulos, University of Nottingham
P-1B-87	End-to-End Joint Semantic Segmentation of Actors and Actions in Video	Jingwei Ji*, Stanford University; Shyamal Buch, Stanford University; Alvaro Soto, Universidad Catolica de Chile; Juan Carlos Niebles, Stanford University
P-1B-88	Learning Discriminative Video Representations Using Adversarial Perturbations	Jue Wang*, ANU; Anoop Cherian, MERL

P-1B-89	DeepWrinkles: Accurate and Realistic Clothing Modeling	Zorah Laehner, TU Munich; Tony Tung*, Facebook / Oculus Research; Daniel Cremers, TUM
P-1B-90	Massively Parallel Video Networks	Viorica Patraucean*, DeepMind; Joao Carreira, DeepMind; Laurent Mazare, DeepMind; Simon Osindero, DeepMind; Andrew Zisserman, University of Oxford

Poster number	Title	Authors
P-2A-01	Unsupervised Person Re-identification by Deep Learning Tracklet Association	Minxian Li*, Nanjing University and Science and Technology; Xiatian Zhu, Queen Mary University, London, UK; Shaogang Gong, Queen Mary University of London
P-2A-02	Instance-level Human Parsing via Part Grouping Network	Ke Gong*, SYSU; Xiaodan Liang, Carnegie Mellon University; Yicheng Li, Sun Yat-sen University; Yimin Chen, sensetime; Liang Lin, Sun Yat-sen University
P-2A-03	Scaling Egocentric Vision: The E-Kitchens Dataset	Dima Damen*, University of Bristol; Hazel Doughty, University of Bristol; Sanja Fidler, University of Toronto; Antonino Furnari, University of Catania; Evangelos Kazakos, University of Bristol; Giovanni Farinella, University of Catania, Italy; Davide Moltisanti, University of Bristol; Jonathan Munro, University of Bristol; Toby Perrett, University of Bristol; Will Price, University of Bristol; Michael Wray, University of Bristol
P-2A-04	Predicting Gaze in Egocentric Video by Learning Task-dependent Attention Transition	Yifei Huang*, The University of Tokyo
P-2A-05	Beyond local reasoning for stereo confidence estimation with deep learning	Fabio Tosi, University of Bologna; Matteo Poggi*, University of Bologna; Antonio Benincasa, University of Bologna; Stefano Mattoccia, University of Bologna
P-2A-06	DeepGUM: Learning Deep Robust Regression with a Gaussian-Uniform Mixture Model	Stéphane Lathuiliere, INRIA; Pablo Mesejo-Santiago, University of Granada; Xavier Alameda-Pineda*, INRIA; Radu Horaud, INRIA
P-2A-07	Into the Twilight Zone: Depth Estimation using Joint Structure-Stereo Optimization	Aashish Sharma*, National University of Singapore; Loong Fah Cheong, NUS

P-2A-08	Generalized Loss-Sensitive Adversarial Learning with Manifold Margins	Marzieh Edraki*, University of Central Florida; Guo-Jun Qi, University of Central Florida
P-2A-09	Adversarial Open Set Domain Adaptation	Kuniaki Saito*, The University of Tokyo; Shohei Yamamoto, The University of Tokyo; Yoshitaka Ushiku, The University of Tokyo; Tatsuya Harada, The University of Tokyo
P-2A-10	Connecting Gaze, Scene and Attention	Eunji Chong*, Georgia Institute of Technology; Nataniel Ruiz, Georgia Institute of Technology; Richard Wang, Georgia Institute of Technology; Yun Zhang, Georgia Institute of Technology
P-2A-11	Multi-modal Cycle-consistent Generalized Zero-Shot Learning	RAFAEL FELIX*, The University of Adelaide; Vijay Kumar B G, University of Adelaide; Ian Reid, University of Adelaide, Australia; Gustavo Carneiro, University of Adelaide
P-2A-12	Understanding Degeneracies and Ambiguities in Attribute Transfer	Attila Szabo*, University of Bern; Qiyang Hu, University of Bern; Tiziano Portenier, University of Bern; Matthias Zwicker, University of Maryland; Paolo Favaro, Bern University, Switzerland
P-2A-13	Start, Follow, Read: End-to-End Full Page Handwriting Recognition	Curtis Wigington*, Brigham Young University; Chris Tensmeyer, Brigham Young University; Brian Davis, Brigham Young University; Bill Barrett, Brigham Young University; Brian Price, Adobe; Scott Cohen, Adobe Research
P-2A-14	Rethinking the Form of Latent States in Image Captioning	Bo Dai*, the Chinese University of Hong Kong; Deming Ye, Tsinghua University; Dahua Lin, The Chinese University of Hong Kong
P-2A-15	ConvNets and ImageNet Beyond Accuracy: Understanding Mistakes and Uncovering Biases	Pierre Stock*, Facebook AI Research; Moustapha Cisse, Facebook AI Research

P-2A-16	Deep Shape Matching	Filip Radenovic*, Visual Recognition Group, CTU Prague; Giorgos Tolias, Vision Recognition Group, Czech Technical University in Prague; Ondrej Chum, Vision Recognition Group, Czech Technical University in Prague
P-2A-17	Neural Stereoscopic Image Style Transfer	Xinyu Gong*, University of Electronic Science and Technology of China; Haozhi Huang, Tencent AI Lab; Lin Ma, Tencent AI Lab; Fumin Shen, UESTC; Wei Liu, Tencent AI Lab; Tong Zhang, Tecent AI Lab
P-2A-18	Semi-supervised FusedGAN for Conditional Image Generation	Navaneeth Bodla*, University of Maryland; Gang Hua, Microsoft Cloud and AI; Rama Chellappa, University of Maryland
P-2A-19	Affine Correspondences between Central Cameras for Rapid Relative Pose Estimation	Iván Eichhardt*, MTA SZTAKI; Mitya Csetverikov, MTA SZTAKI & ELTE
P-2A-20	Bi-directional Feature Pyramid Network with Recursive Attention Residual Modules For Shadow Detection	Lei Zhu*, The Chinese University of Hong Kong; Zijun Deng, South China University of Technology; Xiaowei Hu, The Chinese University of Hong Kong; Chi-Wing Fu, The Chinese University of Hong Kong; Xuemiao Xu, South China University of Technology; Jing Qin, The Hong Kong Polytechnic University; Pheng-Ann Heng, The Chinese Univsersity of Hong Kong
P-2A-21	Joint Learning of Intrinsic Images and Semantic Segmentation	Anil Baslamisli*, University of Amsterdam; Thomas Tiel Groenestege, University of Amsterdam; Partha Das, University of Amsterdam; Hoang-An Le, University of Amsterdam; Sezer Karaoglu, University of Amsterdam; Theo Gevers, University of Amsterdam

P-2A-22	Visual Reasoning with a Multi-hop FiLM Generator	Florian Strub*, University of Lille; Mathieu Seurin, University of Lille; Ethan Perez, Rice University; Harm De Vries, Montreal Institute for Learning Algorithms; Jeremie Mary, Criteo; Philippe Preux, INRIA; Aaron Courville, MILA, Université de Montréal; Olivier Pietquin, GoogleBrain
P-2A-23	View-graph Selection Framework for SfM	Rajvi Shah*, IIIT Hyderabad; Visesh Chari, INRIA; P. J. Narayanan, IIIT-Hyderabad
P-2A-24	Fine-grained Video Categorization with Redundancy Reduction Attention	Chen Zhu, University of Maryland; Xiao Tan, Baidu Inc.; Feng Zhou, Baidu Inc.; Xiao Liu, Baidu Research; Kaiyu Yue*, Baidu Inc.; Errui Ding, Baidu Inc.; Yi Ma, UC Berkeley
P-2A-25	Space-time Knowledge for Unpaired Image-to-Image Translation	Aayush Bansal*, Carnegie Mellon University; Shugao Ma, Facebook / Oculus; Deva Ramanan, Carnegie Mellon University; Yaser Sheikh, CMU
P-2A-26	Integral Human Pose Regression	Xiao Sun*, Microsoft Research Asia; Bin Xiao, MSR Asia; Fangyin Wei, Peking University; Shuang Liang, Tongji University; Yichen Wei, MSR Asia
P-2A-27	Recurrent Tubelet Proposal and Recognition Networks for Action Detection	Dong Li, University of Science and Technology of China; Zhaofan Qiu, University of Science and Technology of China; Qi Dai, Microsoft Research; Ting Yao*, Microsoft Research; Tao Mei, JD.com
P-2A-28	Learning to Predict Crisp Edge	Ruoxi Deng*, Central South University; Chunhua Shen, University of Adelaide; Shengjun Liu, Central South University; Huibing Wang, Dalian University of Technology; Xinru Liu, Central South University
P-2A-29	Open Set Learning with Counterfactual Images	Lawrence Neal*, Oregon State University; Matthew Olson, Oregon State University; Xiaoli Fern, Oregon State University; Weng-Keen Wong, Oregon State University; Fuxin Li, Oregon State University

P-2A-30	Estimating the Success of Unsupervised Image to Image Translation	Lior Wolf, Tel Aviv University, Israel; Sagie Benaim*, Tel Aviv University; Tomer Galanti, Tel Aviv University
P-2A-31	Joint Map and Symmetry Synchronization	Qixing Huang*, The University of Texas at Austin; Xiangru Huang, University of Texas at Austin; Zhenxiao Liang, Tsinghua University; Yifan Sun, The University of Texas at Austin
P-2A-32	Single Image Water Hazard Detection using FCN with Reflection Attention Units	Xiaofeng Han, Nanjing University of Science and Technology; Chuong Nguyen*, CSIRO Data61; Shaodi You, Data61-CSIRO, Australia; Jianfeng Lu, Nanjing University of Science and Technology
P-2A-33	Realtime Time Synchronized Event-based Stereo	Alex Zhu*, University of Pennsylvania; Yibo Chen, University of Pennsylvania; Kostas Daniilidis, University of Pennsylvania
P-2A-34	Transferring GANs: generating images from limited data	yaxing wang*, Computer Vision Center; Chenshen Wu, Computer Vision Center; Luis Herranz, Computer Vision Center (Ph.D.); Joost van de Weijer, Computer Vision Center; Abel Gonzalez-Garcia, Computer Vision Center; BOGDAN RADUCANU, Computer Version Center, Edifici
P-2A-35	To learn image super-resolution, use a GAN to learn how to do image degradation first	Adrian Bulat*, University of Nottingham; Jing Yang, University of Nottingham; Georgios Tzimiropoulos, University of Nottingham
P-2A-36	Unsupervised CNN-based co-saliency detection with graphical optimization	Kuang-Jui Hsu*, Academia Sinica; Chung-Chi Tsai, Texas A&M University; Yen-Yu Lin, Academia Sinica; Xiaoning Qian, Texas A&M University; Yung-Yu Chuang, National Taiwan University
P-2A-37	Fast Light Field Reconstruction With Deep Coarse-To-Fine Modeling of Spatial-Angular Clues	Henry W. F. Yeung, the University of Sydney; Junhui Hou*, City University of Hong Kong, Hong Kong; Jie Chen, Nanyang Technological University; Yuk Ying Chung, the University of Sydney; Xiaoming Chen, University of Science and Technology of China

P-2A-38	Unified Perceptual Parsing for Scene Understanding	Tete Xiao*, Peking University; Yingcheng Liu, Peking University; Yuning Jiang, Megvii(Face++) Inc; Bolei Zhou, MIT; Jian Sun, Megvii, Face++
P-2A-39	PARN: Pyramidal Affine Regression Networks for Dense Semantic Correspondence Estimation	Sangryul Jeon*, Yonsei university; Seungryung Kim, Yonsei University; Dongbo Min, Ewha Womans University; Kwanghoon Sohn , Yonsei Univ.
P-2A-40	Structural Consistency and Controllability for Diverse Colorization	Safa Messaoud*, University of Illinois at Urbana Champaign; Alexander Schwing, UIUC; David Forsyth, Univeristy of Illinois at Urbana-Champaign
P-2A-41	Online Multi-Object Tracking with Dual Matching Attention Networks	Ji Zhu, Shanghai Jiao Tong University; Hua Yang*, Shanghai Jiao Tong University; Nian Liu, Northwestern Polytechnical University; Minyoung Kim, Perceptive Automata; Wenjun Zhang, Shanghai Jiao Tong University; Ming-Hsuan Yang, University of California at Merced
P-2A-42	MaskConnect: Connectivity Learning by Gradient Descent	Karim Ahmed*, Dartmouth College; Lorenzo Torresani, Dartmouth College
P-2A-43	FloorNet: A Unified Framework for Floorplan Reconstruction from 3D Scans	Chen Liu*, Washington University in St. Louis; Jiaye Wu, Washington University in St.Louis; Yasutaka Furukawa, Simon Fraser University
P-2A-44	Image Manipulation with Perceptual Discriminators	Diana Sungatullina*, Skolkovo Institute of Science and Technology; Egor Zakharov, Skolkovo Institute of Science and Technology; Dmitry Ulyanov, Skolkovo Institute of Science and Technology; Victor Lempitsky, Skoltech
P-2A-45	Transductive Centroid Projection for Semi-supervised Large-scale Recognition	Yu Liu*, The Chinese University of Hong Kong; Xiaogang Wang, Chinese University of Hong Kong, Hong Kong; Guanglu Song, Sensetime; Jing Shao, Sensetime

P-2A-46	Eigendecomposition-free Training of Deep Networks with Zero Eigenvalue-based Losses	Zheng Dang*, Xi'an Jiaotong University; Kwang Moo Yi, University of Victoria; Yinlin Hu, EPFL; Fei Wang, Xi'an Jiaotong University; Pascal Fua, EPFL, Switzerland; Mathieu Salzmann, EPFL
P-2A-47	Self-supervised Knowledge Distillation Using Singular Value Decomposition	SEUNG HYUN LEE, Inha University; Daeha Kim, Inha University ; Byung Cheol Song*, Inha University
P-2A-48	Snap Angle Prediction for 360° Panoramas	Bo Xiong*, University of Texas at Austin; Kristen Grauman, University of Texas
P-2A-49	Saliency Preservation in Low-Resolution Grayscale Images	Shivanthan Yohanandan*, RMIT University; Adrian Dyer, RMIT University; Dacheng Tao, University of Sydney; Andy Song, RMIT University
P-2A-50	PPF-FoldNet: Unsupervised Learning of Rotation Invariant 3D Local Descriptors	Tolga Birdal*, TU Munich; Haowen Deng, Technical University of Munich; Slobodan Ilic, Siemens AG
P-2A-51	BusterNet: Detecting Copy-Move Image Forgery with Source/Target Localization	Rex Yue Wu*, USC ISI; Wael Abd-Almageed, Information Sciences Institute; Prem Natarajan, USC ISI
P-2A-52	Double JPEG Detection in Mixed JPEG Quality Factors using Deep Convolutional Neural Network	Jin-Seok Park*, Korea Advanced Institute of Science and Technology (KAIST); Donghyeon Cho, KAIST; Wonhyuk Ahn, KAIST; Heung-Kyu Lee, Korea Advanced Institute of Science and Technology (KAIST)
P-2A-53	Unsupervised holistic image generation from key local patches	Donghoon Lee*, Seoul National University; Sangdoo Yun, Clova AI Research, NAVER Corp.; Sungjoon Choi, Seoul National University; Hwiyeon Yoo, Seoul National University; Ming-Hsuan Yang, University of California at Merced; Songhwai Oh, Seoul National University

P-2A-54	CrossNet: An End-to-end Reference-based Super Resolution Network using Cross-scale Warping	Haitian Zheng, HKUST; Mengqi Ji, HKUST; Haoqian Wang, Tsinghua University; Yebin Liu*, Tsinghua University; Lu Fang, Tsinghua University
P-2A-55	DCAN: Dual Channel-wise Alignment Networks for Unsupervised Scene Adaptation	Zuxuan Wu*, UMD; Xintong Han, University of Maryland, USA; Yen-Liang Lin, GE Global Research ; Gokhan Uzunbas, Avitas Systems-GE Venture; Tom Goldstein, University of Maryland, College Park; Ser-Nam Lim, GE Global Research; Larry Davis, University of Maryland
P-2A-56	YouTube-VOS: Sequence-to-Sequence Video Object Segmentation	Ning Xu*, Adobe Research; Linjie Yang, Snap Research; Dingcheng Yue, UIUC; Jianchao Yang, Snap; Brian Price, Adobe; Jimei Yang, Adobe; Scott Cohen, Adobe Research; Yuchen Fan, Image Formation and Processing (IFP) Group, University of Illinois at Urbana-Champaign; Yuchen Liang, UIUC; Thomas Huang, University of Illinois at Urbana Champaign
P-2A-57	Selfie Video Stabilization	Jiyang Yu*, University of California San Diego; Ravi Ramamoorthi, University of California San Diego
P-2A-58	Videos as Space-Time Region Graphs	Xiaolong Wang*, CMU; Abhinav Gupta, CMU
P-2A-59	Parallel Feature Pyramid Network for Object Detection	Seung-Wook Kim*, Korea University; Hyong-Keun Kook, Korea University; Jee-Young Sun, Korea University; Mun-Cheon Kang, Korea University; Sung-Jea Ko, Korea University
P-2A-60	Goal-Oriented Visual Question Generation via Intermediate Rewards	Junjie Zhang, University of Technology, Sydney; Qi Wu*, University of Adelaide; Chunhua Shen, University of Adelaide; Jian Zhang, UTS; Jianfeng Lu, Nanjing University of Science and Technology; Anton Van Den Hengel, University of Adelaide

P-2A-61	WildDash - Creating Hazard-Aware Benchmarks	Oliver Zendel*, AIT Austrian Institute of Technology; Katrin Honauer, Heidelberg University; Markus Murschitz, AIT Austrian Institute of Technology; Daniel Steininger, AIT Austrian Institute of Technology; Gustavo Fernandez, n/a
P-2A-62	Reinforced Temporal Attention and Split-Rate Transfer for Depth-Based Person Re-identification	Nikolaos Karianakis*, Microsoft; Zicheng Liu, Microsoft; Yinpeng Chen, Microsoft; Stefano Soatto, UCLA
P-2A-63	DF-Net: Unsupervised Joint Learning of Depth and Flow using Cross-Network Consistency	Yuliang Zou*, Virginia Tech; Zelun Luo, Stanford University; Jia-Bin Huang, Virginia Tech
P-2A-64	Generating Multimodal Human Dynamics with a Transformation based Representation	Xinchen Yan*, University of Michigan; Akash Rastogi, UM; Ruben Villegas, University of Michigan; Eli Shechtman, Adobe Research, US; Sunkavalli Kalyan, Adobe Research; Sunil Hadap, Adobe; Ersin Yumer, Argo AI; Honglak Lee, UM
P-2A-65	Learning Rigidity in Dynamic Scenes with a Moving Camera for 3D Motion Field Estimation	Zhaoyang Lv*, GEORGIA TECH; Kihwan Kim, NVIDIA; Alejandro Troccoli, NVIDIA; Deqing Sun, NVIDIA; Kautz Jan, NVIDIA; James Rehg, Georgia Institute of Technology
P-2A-66	Learning Visual Question Answering by Bootstrapping Hard Attention	Mateusz Malinowski*, DeepMind; Carl Doersch, DeepMind; Adam Santoro, DeepMind; Peter Battaglia, DeepMind
P-2A-67	Image Reassembly Combining Deep Learning and Shortest Path Problem	Marie-Morgane Paumard*, ETIS; David Picard, ETIS/LIP6; Hedi Tabia, France
P-2A-68	RESOUND: Towards Action Recognition without Representation Bias	Yingwei Li*, UCSD; Nuno Vasconcelos, UC San Diego; Yi Li, University of California San Diego

P-2A-69	Key-Word-Aware Network for Referring Expression Image Segmentation	Hengcan Shi*, University of Electronic Science and Technology of China; Hongliang Li, University of Electronic Science and Technology of China; Fanman Meng, University of Electronic Science and Technology of China; Qingbo Wu, University of Electronic Science and Technology of China
P-2A-70	Mutual Learning to Adapt for Joint Human Parsing and Pose Estimation	Xuecheng Nie*, NUS; Jiashi Feng, NUS; Shuicheng Yan, Qihoo/360
P-2A-71	Simple Baselines for Human Pose Estimation and Tracking	Bin Xiao*, MSR Asia; Haiping Wu, MSR Asia; Yichen Wei, MSR Asia
P-2A-72	Pose Partition Networks for Multi-Person Pose Estimation	Xuecheng Nie*, NUS; Jiashi Feng, NUS; Junliang Xing, National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences; Shuicheng Yan, Qihoo/360
P-2A-73	Wasserstein Divergence For GANs	Jiqing Wu*, ETH Zurich; Zhiwu Huang, ETH Zurich; Janine Thoma, ETH Zurich; Dinesh Acharya, ETH Zurich; Luc Van Gool, ETH Zurich
P-2A-74	A Segmentation-aware Deep Fusion Network for Compressed Sensing MRI	Zhiwen Fan, Xiamen University; Liyan Sun, Xiamen University; Xinghao Ding*, Xiamen University; Yue Huang, Xiamen University; Congbo Cai, Xiamen University; John Paisley, Columbia University
P-2A-75	Deep Metric Learning with Hierarchical Triplet Loss	Weifeng Ge*, The University of Hong Kong
P-2A-76	Generative Adversarial Network with Spatial Attention for Face Attribute Editing	Gang Zhang*, Institute of Computing Technology, CAS; Meina Kan, Institute of Computing Technology, Chinese Academy of Sciences; Shiguang Shan, Chinese Academy of Sciences; Xilin Chen, China

P-2A-77	Proxy Clouds for Live RGB-D Stream Processing and Consolidation	Adrien Kaiser*, Telecom ParisTech; Jose Alonso Ybanez Zepeda, Ayotle SAS; Tamy Boubekeur, Paris Telecom
P-2A-78	Synthetically Supervised Feature Learning for Scene Text Recognition	Yang Liu*, University of Cambridge; Zhaowen Wang, Adobe Research; Hailin Jin, Adobe Research; Ian Wassell, University of Cambridge
P-2A-79	Scale Aggregation Network for Accurate and Efficient Crowd Counting	Xinkun Cao*, Beijing University of Posts and Telecommunications; Zhipeng Wang, School of Communication and Information Engineering, Beijing University of Posts and Telecommunications; Yanyun Zhao, Beijing Univiersity of Posts and Telecommunications; Fei Su, Beijing University of Posts and Telecommunications
P-2A-80	PM-GANs: Discriminative Representation Learning for Action Recognition Using Partial-modalities	Lan Wang, Chongqing Key Laboratory of Signal and Information Processing, Chongqing University of Posts and Telecommunications; Chenqiang Gao*, Chongqing Key Laboratory of Signal and Information Processing, Chongqing University of Posts and Telecommunications; Luyu Yang, Chongqing Key Laboratory of Signal and Information Processing, Chongqing University of Posts and Telecommunications; Yue Zhao, Chongqing Key Laboratory of Signal and Information Processing, Chongqing University of Posts and Telecommunications; Wangmeng Zuo, Harbin Institute of Technology, China; Deyu Meng, Xi'an Jiaotong University
P-2A-81	OmniDepth: Dense Depth Estimation for Indoors Spherical Panoramas.	NIKOLAOS ZIOULIS*, CERTH / CENTRE FOR RESEARCH AND TECHNOLOGY HELLAS; Antonis Karakottas, CERTH / CENTRE FOR RESEARCH AND TECHNOLOGY HELLAS; Dimitrios Zarpalas, CERTH / CENTRE FOR RESEARCH AND TECHNOLOGY HELLAS; Petros Daras, ITI-CERTH, Greece

P-2A-82	Hashing with Binary Matrix Pursuit	Fatih Cakir*, Boston University; Kun He, Boston University; Stan Sclaroff, Boston University
P-2A-83	Probabilistic Video Generation using Holistic Attribute Control	Jiawei He*, Simon Fraser University; Andreas Lehrmann, Facebook; Joe Marino, California Institute of Technology; Greg Mori, Simon Fraser University; Leonid Sigal, University of British Columbia
P-2A-84	Transductive Semi-Supervised Deep Learning using Min-Max Features	Weiwei Shi*, Xi'an Jiaotong University; Yihong Gong, Xi'an Jiaotong University; Chris Ding, UNIVERSITY OF TEXAS AT ARLINGTON; Zhiheng Ma, Xi'an Jiaotong University; Xiaoyu Tao, Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University.; Nanning Zheng, Xi'an Jiaotong University
P-2A-85	Deep Feature Pyramid Reconfiguration for Object Detection	Tao Kong*, Tsinghua; Fuchun Sun, Tsinghua; Wenbing Huang, Tencent AI Lab
P-2A-86	Quadtree Convolutional Neural Networks	Pradeep Kumar Jayaraman*, Nanyang Technological University; Jianhan Mei, Nanyang Technological University; Jianfei Cai, Nanyang Technological University; Jianmin Zheng, Nanyang Technological University
P-2A-87	Correcting the Triplet Selection Bias for Triplet Loss	Baosheng Yu*, The University of Sydney; Tongliang Liu, The University of Sydney; Mingming Gong, CMU & U Pitt; Changxing Ding, South China University of Technology; Dacheng Tao, University of Sydney
P-2A-88	Adversarial Geometry-Aware Human Motion Prediction	Liangyan Gui*, Carnegie Mellon University; XIAODAN LIANG, Carnegie Mellon University; Yu-Xiong Wang, Carnegie Mellon University; José M. F. Moura, Carnegie Mellon University

P-2B-01	3D Motion Sensing from 4D Light Field Gradients	Sizhuo Ma*, University of Wisconsin-Madison; Brandon Smith, University of Wisconsin-Madison; Mohit Gupta, University of Wisconsin-Madison, USA
P-2B-02	A Trilateral Weighted Sparse Coding Scheme for Real-World Image Denoising	XU JUN, The Hong Kong Polytechnic University; Lei Zhang*, Hong Kong Polytechnic University, Hong Kong, China; D. Zhang, The Hong Kong Polytechnic University
P-2B-03	Saliency Detection in 360° Videos	Ziheng Zhang, Shanghaitech University; Yanyu Xu*, Shanghaitech University; Shenghua Gao, Shanghaitech University; Jingyi Yu, Shanghai Tech University
P-2B-04	Learning to Blend Photos	Wei-Chih Hung*, University of California, Merced; Jianming Zhang, Adobe Research; Xiaohui Shen, Adobe Research; Zhe Lin, Adobe Research; Joon-Young Lee, Adobe Research; Ming-Hsuan Yang, University of California at Merced
P-2B-05	Escaping from Collapsing Modes in a Constrained Space	Chieh Lin, National Tsing Hua University; Chia-Che Chang, National Tsing Hua University; Che-Rung Lee, National Tsing Hua University; Hwann-Tzong Chen*, National Tsing Hua University
P-2B-06	Verisimilar Image Synthesis for Accurate Detection and Recognition of Texts in Scenes	Fangneng Zhan, Nanyang Technological University; Shijian Lu*, Nanyang Technological University; Chuhui Xue, Nanyang Technological University
P-2B-07	Layer-structured 3D Scene Inference via View Synthesis	Shubham Tulsiani*, UC Berkeley; Richard Tucker, Google; Noah Snavely, -

P-2B-08	Perturbation Robust Representations of Topological Persistence Diagrams	Anirudh Som*, Arizona State University; Kowshik Thopalli, Arizona State University; Karthikeyan Natesan Ramamurthy, IBM Research; Vinay Venkataraman, Arizona State University; Ankita Shukla, Indraprastha Institute of Information Technology - Delhi; Pavan Turaga, Arizona State University
P-2B-09	Analyzing Clothing Layer Deformation Statistics of 3D Human Motions	Jinlong YANG*, Inria; Jean-Sebastien Franco, INRIA; Franck Hétroy-Wheeler, University of Strasbourg; Stefanie Wuhrer, Inria
P-2B-10	Neural Nonlinear least Squares with Application to Dense Tracking and Mapping	Ronald Clark*, Imperial College London; Michael Bloesch, Imperial; Jan Czarnowski, Imperial College London; Andrew Davison, Imperial College London; Stefan Leutenegger, Imperial College London
P-2B-11	Propagating LSTM: 3D Pose Estimation based on Joint Interdependency	Kyoungoh Lee*, Yonsei University; Inwoong Lee, Yonsei University; Sanghoon Lee, Yonsei University, Korea
P-2B-12	Proximal Dehaze-Net: A Prior Learning-Based Deep Network for Single Image Dehazing	Dong Yang, Xi'an Jiaotong University; JIAN SUN*, Xi'an Jiaotong University
P-2B-13	Attend and Rectify: a gated attention mechanism for fine-grained recovery	Pau Rodriguez Lopez*, Computer Vision Center, Universitat Autonoma de Barcelona; Guillem Cucurull, Computer Vision Center, Universitat Autonoma de Barcelona; Josep Gonfaus, Computer Vision Center; Jordi Gonzalez, UA Barcelona; Xavier Roca, Computer Vision Center, Universitat Autonoma de Barcelona
P-2B-14	Learning to Capture Light Fields through A Coded Aperture Camera	Yasutaka Inagaki*, Nagoya University; Yuto Kobayashi, Nagoya University; Keita Takahashi, Nagoya University; Toshiaki Fujii, Nagoya University; Hajime Nagahara, Osaka University

P-2B-15	AMC: Automated Model Compression and Acceleration with Reinforcement Learning	Yihui He, Xi'an Jiaotong University; Ji Lin, Tsinghua University; Song Han*, MIT
P-2B-16	Extreme Network Compression via Filter Group Approximation	Bo Peng*, Hikvision Research Institute; Wenming Tan, Hikvision Research Institute; Zheyang Li, Hikvision Research Institute; Shun Zhang, Hikvision Research Institute; Di Xie, Hikvision Research Institute; Shiliang Pu, Hikvision Research Institute
P-2B-17	Retrospective Encoders for Video Summarization	Ke Zhang*, USC; Kristen Grauman, University of Texas; Fei Sha, USC
P-2B-18	Optimized Quantization for Highly Accurate and Compact DNNs	Dongqing Zhang, Microsoft Research; Jiaolong Yang*, Microsoft Research Asia (MSRA); Dongqiangzi Ye, Microsoft Research; Gang Hua, Microsoft Cloud and AI
P-2B-19	Universal Sketch Perceptual Grouping	Ke LI*, Queen Mary University of London; Kaiyue Pang, Queen Mary University of London; Jifei Song, Queen Mary, University of London; Yi-Zhe Song, Queen Mary University of London; Tao Xiang, Queen Mary, University of London, UK; Timothy Hospedales, Edinburgh University; Honggang Zhang, Beijing University of Posts and Telecommunications
P-2B-20	Uncertainty Estimates and Multi-Hypotheses Networks for Optical Flow	Eddy Ilg*, University of Freiburg; Özgün Çiçek, University of Freiburg; Silvio Galesso, University of Freiburg; Aaron Klein, Universität Freiburg; Osama Makansi, University of Freiburg; Frank Hutter, University of Freiburg; Thomas Brox, University of Freiburg

P-2B-21	Learning 3D Keypoint Descriptors for Non-Rigid Shape Matching	Hanyu Wang, NLPR, Institute of Automation, Chinese Academy of Sciences; Jianwei Guo*, NLPR, Institute of Automation, Chinese Academy of Sciences; Yan Dong-Ming, NLPR, CASIA; Weize Quan, NLPR, Institute of Automation, Chinese Academy of Sciences; Xiaopeng Zhang, Institute of Automation, Chinese Academy of Sciences
P-2B-22	A Joint Sequence Fusion Model for Video Question Answering and Retrieval	Youngjae Yu, Seoul National University Vision and Learning Lab; Jongseok Kim, Seoul National University Vision and Learning Lab; Gunhee Kim*, Seoul National University
P-2B-23	Deformable Pose Traversal Convolution for 3D Action and Gesture Recognition	Junwu Weng*, Nanyang Technological University; Mengyuan Liu, Nanyang Technological University; Xudong Jiang, Nanyang Technological University; Junsong Yuan, State University of New York at Buffalo, USA
P-2B-24	Fine-Grained Visual Categorization using Meta-Learning Optimization with Sample Selection of Auxiliary Data	Yabin Zhang, South China University of Technology; Tang Hui, South China University of Technology; Kui Jia*, South China University of Technology
P-2B-25	Stereo relative pose from line and point feature triplets	Alexander Vakhitov*, Skoltech; Victor Lempitsky, Skoltech; Yinqiang Zheng, National Institute of Informatics
P-2B-26	Convolutional Block Attention Module	Sanghyun Woo*, KAIST; Jongchan Park, KAIST; Joon-Young Lee, Adobe Research; In So Kweon, KAIST
P-2B-27	EC-Net: an Edge-aware Point set Consolidation Network	Lequan Yu*, The Chinese University of Hong Kong; Xianzhi Li, The Chinese University of Hong Kong; Chi-Wing Fu, The Chinese University of Hong Kong; Danny Cohen-Or, Tel Aviv University; Pheng-Ann Heng, The Chinese Univsersity of Hong Kong

P-2B-28	Video Compression through Image Interpolation	Chao-Yuan Wu*, UT Austin; Nayan Singhal, UT Austin; Philipp Kraehenbuehl, UT Austin
P-2B-29	Burst Image Deblurring Using Permutation Invariant Convolutional Neural Networks	Miika Aittala*, MIT; Fredo Durand, MIT
P-2B-30	HybridNet: Classification and Reconstruction Cooperation for Semi-Supervised Learning	Thomas Robert*, LIP6 / Sorbonne Universite; Nicolas Thome, CNAM, Paris; Matthieu Cord, Sorbonne University
P-2B-31	Structure-from-Motion-Aware PatchMatch for Adaptive Optical Flow Estimation	Daniel Maurer*, University of Stuttgart; Nico Marniok, Universität Konstanz; Bastian Goldluecke, University of Konstanz; Andrés Bruhn, University of Stuttgart
P-2B-32	Joint & Progressive Learning from High-Dimensional Data for Multi-Label Classification	Danfeng Hong*, Technical University of Munich (TUM); German Aerospace Center (DLR); Naoto Yokoya, RIKEN Center for Advanced Intelligence Project (AIP); Jian Xu, German Aerospace Center (DLR); Xiaoxiang Zhu, DLR&-TUM
P-2B-33	SDC-Net: Video prediction using spatially-displaced convolution	Fitsum Reda*, NVIDIA; Guilin Liu, NVIDIA; Kevin Shih, NVIDIA; Robert Kirby, Nvidia; Jon Barker, Nvidia; David Tarjan, Nvidia; Andrew Tao, NVIDIA; Bryan Catanzaro, NVIDIA
P-2B-34	Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation	Liang-Chieh Chen*, Google Inc.; Yukun Zhu, Google Inc.; George Papandreou, Google; Florian Schroff, Google Inc.; Hartwig Adam, Google
P-2B-35	VQA-E: Explaining, Elaborating, and Enhancing Your Answers for Visual Questions	Qing Li*, University of Science and Technology of China; Qingyi Tao, Nanyang Techonological University; Shafiq Joty, Nanyang Technological University; Jianfei Cai, Nanyang Technological University; Jiebo Luo, U. Rochester

P-2B-36	Image Super-Resolution Using Very Deep Residual Channel Attention Networks	Yulun Zhang*, Northeastern University; Kunpeng Li, Northeastern University; kai li, northeastern university; Lichen Wang, Northeastern University; Bineng Zhong, Huaqiao University; YUN FU, Northeastern University
P-2B-37	Urban Zoning Using Higher-Order Markov Random Fields on Multi-View Imagery Data	Tian Feng, University of New South Wales; Quang-Trung Truong, SUTD; Thanh Nguyen*, Deakin University, Australia; Jing Yu Koh, SUTD; Lap-Fai Yu, UMass Boston; Sai-Kit Yeung, Singapore University of Technology and Design; Alexander Binder, Singapore University of Technology and Design
P-2B-38	Clustering Kernels for Compressing the Convolutional Neural Networks	Sanghyun Son, Seoul National University; Seungjun Nah, Seoul National University; Kyoung Mu Lee*, Seoul National University
P-2B-39	Explainable Neural Computation via Stack Neural Module Networks	Ronghang Hu*, University of California, Berkeley; Jacob Andreas, UC Berkeley; Trevor Darrell, UC Berkeley; Kate Saenko, Boston University
P-2B-40	Quaternion Convolutional Neural Networks	Xuanyu Zhu*, Shanghai Jiao Tong University; Yi Xu, Shanghai Jiao Tong University; Hongteng Xu, Duke University; Changjian Chen, Shanghai Jiao Tong University
P-2B-41	Lip Movements Generation at a Glance	Lele Chen*, University of Rochester; Zhiheng Li, WuHan University; Ross Maddox, University of Rochester; Zhiyao Duan, Unversity of Rochester; Chenliang Xu, University of Rochester
P-2B-42	Toward Scale-Invariance and Position-Sensitive Object Proposal Networks	Hsueh-Fu Lu, Umbo Computer Vision; Ping-Lin Chang*, Umbo Computer Vision; Xiaofei Du, Umbo Computer Vision
P-2B-43	Constraints Matter in Deep Neural Network Compression	Changan Chen, Simon Fraser University; Fred Tung*, Simon Fraser University; Naveen Vedula, Simon Fraser University; Greg Mori, Simon Fraser University

P-2B-44	MRF Optimization with Separable Convex Prior on Partially Ordered Labels	Csaba Domokos*, Technical University of Munich; Frank Schmidt, BCAI; Daniel Cremers, TUM
P-2B-45	Switchable Temporal Propagation Network	Sifei Liu*, NVIDIA; Ming-Hsuan Yang, University of California at Merced; Guangyu Zhong, Dalian University of Technology; Jinwei Gu, Nvidia; Shalini De Mello, NVIDIA Research; Kautz Jan, NVIDIA; Varun Jampani, Nvidia Research
P-2B-46	T2Net: Synthetic-to-Realistic Translation for Solving Single-Image Depth Estimation Tasks	Chuanxia Zheng*, Nanyang Technological University; Tat-Jen Cham, Nanyang Technological University; Jianfei Cai, Nanyang Technological University
P-2B-47	ArticulatedFusion: Real-time Reconstruction of Motion, Geometry and Segmentation Using a Single Depth Camera	Chao Li*, The University of Texas at Dallas; Zheheng Zhao, The University of Texas at Dallas; Xiaohu Guo, The University of Texas at Dallas
P-2B-48	NNEval: Neural Network based Evaluation Metric for Image Captioning	Naeha Sharif*, University of Western Australia; Lyndon White, University of Western Australia; Mohammed Bennamoun, University of Western Australia; Syed Afaq Ali Shah, Department of Computer Science and Software Engineering, The University of Western Australia
P-2B-49	Coreset-Based Convolutional Neural Network Compression	Abhimanyu Dubey*, Massachusetts Institute of Technology; Moitreya Chatterjee, University of Illinois at Urbana Champaign; Ramesh Raskar, Massachusetts Institute of Technology; Narendra Ahuja, University of Illinois at Urbana-Champaign, USA
P-2B-50	Context Refinement for Object Detection	Zhe Chen*, University of Sydney; Shaoli Huang, University of Sydney; Dacheng Tao, University of Sydney

P-2B-51	Real-time Actor-Critic-Tracking	Boyu Chen*, Dalian University of Technology; Dong Wang, Dalian University of Technology; Peixia Li, Dalian University of Technology; Huchuan Lu, Dalian University of Technology
P-2B-52	Partial Adversarial Domain Adaptation	Zhangjie Cao, Tsinghua University; Lijia Ma, Tsinghua University; Mingsheng Long*, Tsinghua University; Jianmin Wang, Tsinghua University, China
P-2B-53	Localization Recall Precision (LRP): A New Performance Metric for Object Detection	Kemal Oksuz*, Middle East Technical University; Bar?? Can Çam, Roketsan; Emre Akbas, Middle East Technical University; Sinan Kalkan, Middle East Technical University
P-2B-54	Improving Embedding Generalization via Scalable Neighborhood Component Analysis	Zhirong Wu*, UC Berkeley; Alexei Efros, UC Berkeley; Stella Yu, UC Berkeley / ICSI
P-2B-55	Leveraging Motion Priors in Videos for Improving Human Segmentation	Yu-Ting Chen*, NTHU; Wen-Yen Chang, NTHU; Hai-Lun Lu, NTHU; Tingfan Wu, Umbo Computer Vision; Min Sun, NTHU
P-2B-56	Recurrent Squeeze-and-Excitation Context Aggregation Net for Single Image Deraining	Xia Li*, Peking University Shenzhen Graduate School; Jianlong Wu, Peking University; Zhouchen Lin, Peking University; Hong Liu, Peking University Shenzhen Graduate School; Hongbin Zha, Peking University, China
P-2B-57	Statistically-motivated Second-order Pooling	Kaicheng Yu*, EPFL; Mathieu Salzmann, EPFL
P-2B-58	SegStereo: Exploiting Semantic Information for Disparity Estimation	Guorun Yang*, Tsinghua University; Hengshuang Zhao, The Chinese University of Hong Kong; Jianping Shi, Sensetime Group Limited; Jia Jiaya, Chinese University of Hong Kong

P-2B-59	Small-scale Pedestrian Detection Based on Somatic Topology Localization and Temporal Feature Aggregation	Tao Song, Hikvision Research Institute; Leiyu Sun, Hikvision Research Institute; Di Xie*, Hikvision Research Institute; Haiming Sun, Hikvision Research Institute; Shiliang Pu, Hikvision Research Institute
P-2B-60	Object Detection with an Aligned Spatial-Temporal Memory	Fanyi Xiao*, University of California Davis; Yong Jae Lee, University of California, Davis
P-2B-61	Learning to Drive with 360° Surround-View Cameras and a Map	Simon Hecker*, ETH Zurich; Dengxin Dai, ETH Zurich; Luc Van Gool, ETH Zurich
P-2B-62	Monocular Scene Parsing and Reconstruction using 3D Holistic Scene Grammar	Siyuan Huang*, UCLA; Siyuan Qi, UCLA; Yixin Zhu, UCLA; Yinxue Xiao, University of California, Los Angeles; Yuanlu Xu, University of California, Los Angeles; Song-Chun Zhu, UCLA
P-2B-63	Coded Illumination and Imaging for Fluorescence Based Classification	Yuta Asano, Tokyo Institute of Technology; Misaki Meguro, Tokyo Institute of Technology; Chao Wang, Kyushu Institute of Technology; Antony Lam*, Saitama University; Yinqiang Zheng, National Institute of Informatics; Takahiro Okabe, Kyushu Institute of Technology; Imari Sato, National Institute of Informatics
P-2B-64	Modality Distillation with Multiple Stream Networks for Action Recognition	Nuno Garcia, IIT; Pietro Morerio*, IIT; Vittorio Murino, Istituto Italiano di Tecnologia
P-2B-65	VideoMatch: Matching based Video Object Segmentation	Yuan-Ting Hu*, University of Illinois at Urbana-Champaign; Jia-Bin Huang, Virginia Tech; Alexander Schwing, UIUC
P-2B-66	Superpixel Sampling Networks	Varun Jampani*, Nvidia Research; Deqing Sun, NVIDIA; Ming-Yu Liu, NVIDIA; Ming-Hsuan Yang, University of California at Merced; Kautz Jan, NVIDIA

P-2B-67	Deep Bilinear Learning for RGB-D Action Recognition	HU Jian-Fang, Sun Yat-sen University; Jason Wei Shi Zheng*, Sun Yat Sen University; Pan Jiahui, Sun Yat-sen University; Jian-Huang Lai, Sun Yat-sen University; Jianguo Zhang, University of Dundee
P-2B-68	Multi-object Tracking with Neural Gating using bilinear LSTMs	Chanho Kim*, Georgia Tech; Fuxin Li, Oregon State University; James Rehg, Georgia Institute of Technology
P-2B-69	Direct Sparse Odometry With Rolling Shutter	David Schubert*, Technical University of Munich; Vladyslav Usenko, TU Munich; Nikolaus Demmel, TUM; Joerg Stueckler, Technical University of Munich; Daniel Cremers, TUM
P-2B-70	Person Search via A Mask-guided Two-stream CNN Model	Di Chen*, Nanjing University of Science and Techonology; Shanshan Zhang, Max Planck Institute for Informatics; Wanli Ouyang, CUHK; Jian Yang, Nanjing University of Science and Technology; Ying Tai, Tencent
P-2B-71	Imagine This! Scripts to Compositions to Videos	Tanmay Gupta*, UIUC; Dustin Schwenk, Allen Institute for Artificial Intelligence; Ali Farhadi, University of Washington; Derek Hoiem, University of Illinois at Urbana-Champaign; Aniruddha Kembhavi, Allen Institute for Artificial Intelligence
P-2B-72	Multiresolution Tree Networks for Point Cloud Procesing	Matheus Gadelha*, University of Massachusetts Amherst; Subhransu Maji, University of Massachusetts, Amherst; Rui Wang, U Massachusetts
P-2B-73	Quantization Mimic: Towards Very Tiny CNN for Object Detection	Yi Wei*, Tsinghua University; Xinyu Pan, MMLAB, CUHK; Hongwei Qin, SenseTime; Junjie Yan, Sensetime; Wanli Ouyang, CUHK
P-2B-74	Multi-scale Residual Network for Image Super-Resolution	Juncheng Li, East China Normal University; Faming Fang*, East China Normal University; Kangfu Mei, Jiangxi Normal University; Guixu Zhang, East China Normal University

P-2B-75	BodyNet: Volumetric Inference of 3D Human Body Shapes	Gul Varol*, INRIA; Duygu Ceylan, Adobe Research; Bryan Russell, Adobe Research; Jimei Yang, Adobe; Ersin Yumer, Argo AI; Ivan Laptev, INRIA Paris; Cordelia Schmid, INRIA
P-2B-76	3D Recurrent Neural Networks with Context Fusion for Point Cloud Semantic Segmentation	Xiaoqing Ye*, SIMIT; Jiamao Li, SIMIT; Hexiao Huang, Shanghai Opening University; Xiaolin Zhang, SIMIT
P-2B-77	Robust Anchor Embedding for Unsupervised Video Re-Identification in the Wild	Mang YE*, Hong Kong Baptist University; Xiangyuan Lan, Department of Computer Science, Hong Kong Baptist University; PongChi Yuen, Department of Computer Science, Hong Kong Baptist University
P-2B-78	Towards Robust Neural Networks via Random Self-ensemble	Xuanqing Liu, UC Davis Department of Computer Science; Minhao Cheng, University of California, Davis; Huan Zhang, UC Davis; Cho-Jui Hsieh*, UC Davis Department of Computer Science and Statistics
P-2B-79	SpiderCNN: Deep Learning on Point Sets with Parameterized Convolutional Filters	Yifan Xu, Tsinghua University; Tianqi Fan, Multimedia Laboratory, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences; Mingye Xu, Multimedia Laboratory, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences; Long Zeng, Tsinghua University; Yu Qiao*, Multimedia Laboratory, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences
P-2B-80	CIRL: Controllable Imitative Reinforcement Learning for Vision-based Self-driving	Xiaodan Liang*, Carnegie Mellon University; Tairui Wang, Petuum Inc; Luona Yang, Carnegie Mellon University; Eric Xing, Petuum Inc.
P-2B-81	Normalized Blind Deconvolution	Meiguang Jin*, University of Bern; Stefan Roth, TU Darmstadt; Paolo Favaro, Bern University, Switzerland

P-2B-82	Few-Shot Human Motion Prediction via Meta-Learning	Liangyan Gui*, Carnegie Mellon University; Yu-Xiong Wang, Carnegie Mellon University; José M. F. Moura, Carnegie Mellon University
P-2B-83	Learning to Segment via Cut-and-Paste	Tal Remez*, Tel-Aviv University; Matthew Brown, Google; Jonathan Huang, Google
P-2B-84	Weakly-supervised 3D Hand Pose Estimation from Monocular RGB Images	Yujun Cai*, Nanyang Technological University; Liuhao Ge, NTU; Jianfei Cai, Nanyang Technological University; Junsong Yuan, State University of New York at Buffalo, USA
P-2B-85	DeepIM: Deep Iterative Matching for 6D Pose Estimation	Yi Li*, Tsinghua University; Gu Wang, Tsinghua University; Xiangyang Ji, Tsinghua University; Yu Xiang, University of Michigan; Dieter Fox, University of Washington
P-2B-86	Jointly Discovering Visual Objects and Spoken Words from Raw Sensory Input	David Harwath*, MIT CSAIL; Adria Recasens, Massachusetts Institute of Technology; Dídac Surís, Universitat Politecnica de Catalunya; Galen Chuang, MIT; Antonio Torralba, MIT; James Glass, MIT
P-2B-87	A Style-aware Content Loss for Real-time HD Style Transfer	Artsiom Sanakoyeu*, Heidelberg University; Dmytro Kotovenko, Heidelberg University; Bjorn Ommer, Heidelberg University
P-2B-88	Implicit 3D Orientation Learning for 6D Object Detection from RGB Images	Martin Sundermeyer*, German Aerospace Center (DLR); Zoltan Marton, DLR; Maximilian Durner, DLR; Rudolph Triebel, German Aerospace Center (DLR)
P-2B-89	Scale-Awareness of Light Field Camera based Visual Odometry	Niclas Zeller*, Karlsruhe University of Applied Sciences; Franz Quint, Karlsruhe University of Applied Sciences; Uwe Stilla, Technische Universitaet Muenchen
P-2B-90	Audio-Visual Scene Analysis with Self-Supervised Multisensory Features	Andrew Owens*, UC Berkeley; Alexei Efros, UC Berkeley

Poster-number	Title	Authors
P-3A-01	Efficient Sliding Window Computation for NN-Based Template Matching	Lior Talker*, Haifa University; Yael Moses, IDC, Israel; Ilan Shimshoni, University of Haifa
P-3A-02	Active Stereo Net: End-to-End Self-Supervised Learning for Active Stereo Systems	Yinda Zhang*, Princeton University; Sean Fanello, Google; Sameh Khamis, Google; Christoph Rhemann, Google; Julien Valentin, Google; Adarsh Kowdle, Google; Vladimir Tankovich, Google; Shahram Izadi, Google; Thomas Funkhouser, Princeton, USA
P-3A-03	GAL: Geometric Adversarial Loss for Single-View 3D-Object Reconstruction	Li Jiang*, The Chinese University of Hong Kong; Xiaojuan Qi, CUHK; Shaoshuai SHI, The Chinese University of Hong Kong; Jia Jiaya, Chinese University of Hong Kong
P-3A-04	Learning to Reconstruct High-quality 3D Shapes with Cascaded Fully Convolutional Networks	Yan-Pei Cao*, Tsinghua University; Zheng-Ning Liu, Tsinghua University; Zheng-Fei Kuang, Tsinghua University; Shi-Min Hu, Tsinghua University
P-3A-05	Deep Reinforcement Learning with Iterative Shift for Visual Tracking	Liangliang Ren, Tsinghua University; Xin Yuan, Tsinghua University; Jiwen Lu*, Tsinghua University; Ming Yang, Horizon Robotics; Jie Zhou, Tsinghua University, China
P-3A-06	Enhancing Image Geo-localization by Combinatorial Partitioning of Maps	Paul Hongsuck Seo*, POSTECH; Tobias Weyand, Google Inc.; Jack Sim, Google LLC; Bohyung Han, Seoul National University
P-3A-07	Bayesian Instance Segmentation in Open Set World	Trung Pham*, NVIDIA; Vijay Kumar B G, University of Adelaide; Thanh-Toan Do, The University of Adelaide; Gustavo Carneiro, University of Adelaide; Ian Reid, University of Adelaide, Australia

P-3A-08	Characterize Adversarial Examples Based on Spatial Consistency Information for Semantic Segmentation	CHAOWEI XIAO, University of Michigan, Ann Arbor; Ruizhi Deng, Simon Fraser University; Bo Li*, University of Illinois at Urbana Champaign; Fisher Yu, UC Berkeley; mingyan liu, university of Michigan, Ann Arbor; Dawn Song, UC Berkeley
P-3A-09	CubeNet: Equivariance to 3D Rotation and Translation	Daniel Worrall*, UCL; Gabriel Brostow, University College London
P-3A-10	3D Face Reconstruction from Light Field Images: A Model-free Approach	Mingtao Feng, Hunan Unversity; Syed Zulqarnain Gilani*, The University of Western Australia; Yaonan Wang, Hunan University; Ajmal Mian, University of Western Australia
P-3A-11	stagNet: An Attentive Semantic RNN for Group Activity Recognition	Mengshi Qi*, Beihang University; Jie Qin, ETH Zurich; Annan Li, Beijing University of Aeronautics and Astronautics; Yunhong Wang, State Key Laboratory of Virtual Reality Technology and System, Beihang University, Beijing 100191, China; Jiebo Luo, U. Rochester; Luc Van Gool, ETH Zurich
P-3A-12	Supervising the new with the old: learning SFM from SFM	Maria Klodt*, University of Oxford; Andrea Vedaldi, Oxford University
P-3A-13	PSANet: Point-wise Spatial Attention Network for Scene Parsing	Hengshuang Zhao*, The Chinese University of Hong Kong; Yi ZHANG, The Chinese University of Hong Kong; Shu Liu, CUHK; Jianping Shi, Sensetime Group Limited; Chen Change Loy, Chinese University of Hong Kong; Dahua Lin, The Chinese University of Hong Kong; Jia Jiaya, Chinese University of Hong Kong
P-3A-14	FishEyeRecNet: A Multi-Context Collaborative Deep Network for Fisheye Image Recti_x000c_cation	Xiaoqing Yin*, University of Sydney; Xinchao Wang, Stevens Institute of Technology; Jun Yu, HDU; Maojun Zhang, National University of Defense Technology, China; Pascal Fua, EPFL, Switzerland; Dacheng Tao, University of Sydney

P-3A-15	ELEGANT: Exchanging Latent Encodings with GAN for Transferring Multiple Face Attributes	Taihong Xiao*, Peking University; Jiapeng Hong, Peking University; Jinwen Ma, Peking University
P-3A-16	Deep Bilevel Learning	Simon Jenni*, Universität Bern; Paolo Favaro, Bern University, Switzerland
P-3A-17	ADVIO: An Authentic Dataset for Visual-Inertial Odometry	Santiago Cortes, Aalto University; Arno Solin*, Aalto University; Esa Rahtu, Tampere University of Technology; Juho Kannala, Aalto University, Finland
P-3A-18	D2S: Densely Segmented Supermarket Dataset	Patrick Follmann*, MVTec Software GmbH; Tobias Böttger, MVTec Software GmbH; Philipp Härtinger, MVTec Software GmbH; Rebecca König, MVTec Software GmbH
P-3A-19	PyramidBox: A Context-assisted Single Shot Face Detector	Xu Tang, Baidu; Daniel Du*, Baidu; Zeqiang He, Baidu; jingtuo liu, baidu
P-3A-20	Structured Siamese Network for Real-Time Visual Tracking	Yunhua Zhang, Dalian University of Technology; Lijun Wang, Dalian University of Technology; Dong Wang, Dalian University of Technology; Mengyang Feng, Dalian University of Technology; Huchuan Lu*, Dalian University of Technology; Jinqing Qi, Dalian University of Technology
P-3A-21	Probabilistic Signed Distance Function for On-the-fly Scene Reconstruction	Wei Dong*, Peking University; Qiuyuan Wang, Peking University; Xin Wang, Peking University; Hongbin Zha, Peking University, China
P-3A-22	3D Vehicle Trajectory Reconstruction in Monocular Video Data Using Environment Structure Constraints	Sebastian Bullinger*, Fraunhofer IOSB; Christoph Bodensteiner, Fraunhofer IOSB; Michael Arens, Fraunhofer IOSB; Rainer Stiefelhagen, Karlsruhe Institute of Technology

P-3A-23	Unsupervised Image-to-Image Translation with Stacked Cycle-Consistent Adversarial Networks	Minjun Li*, Fudan University; Haozhi Huang, Tencent AI Lab; Lin Ma, Tencent AI Lab; Wei Liu, Tencent AI Lab; Tong Zhang, Tecent AI Lab; Yu-Gang Jiang, Fudan University
P-3A-24	Pose-Normalized Image Generation for Person Re-identification	Xuelin Qian, Fudan University; Yanwei Fu*, Fudan Univ.; Tao Xiang, Queen Mary, University of London, UK; Wenxuan Wang, Fudan University; Jie Qiu, Nara Institute of Science and Technology; Yang Wu, Nara Institute of Science and Technology; Yu-Gang Jiang, Fudan University; Xiangyang Xue, Fudan University
P-3A-25	Action Anticipation with RBF Kernelized Feature Mapping RNN	Yuge Shi*, Australian National University; Basura Fernando, Australian National University; RICHARD HARTLEY, Australian National University, Australia
P-3A-26	Rendering Portraitures from Monocular Camera and Beyond	Xiangyu Xu*, Tsinghua University; Deqing Sun, NVIDIA; Sifei Liu, NVIDIA; Wenqi Ren, Institute of Information Engineering, Chinese Academy of Sciences; Yu-Jin Zhang, Tsinghua University; Ming-Hsuan Yang, University of California at Merced; Jian Sun, Megvii, Face++
P-3A-27	Recovering 3D Planes from a Single Image via Convolutional Neural Networks	Fengting Yang*, Pennsylvania State University ; Zihan Zhou, Penn State University
P-3A-28	The Devil of Face Recognition is in the Noise	Liren Chen*, Sensetime Group Limited; Fei Wang, SenseTime; Cheng Li, SenseTime Research; Shiyao Huang, SenseTime Co Ltd; Yanjie Chen, sensetime; Chen Qian, SenseTime; Chen Change Loy, Chinese University of Hong Kong
P-3A-29	3DMV: Joint 3D-Multi-View Prediction for 3D Semantic Scene Segmentation	Angela Dai*, Stanford University; Matthias Niessner, Technical University of Munich

P-3A-30	Joint optimization for compressive video sensing and reconstruction under hardware constraints	Michitaka Yoshida*, Kyushu University; Akihiko Torii, Tokyo Institute of Technology, Japan; Masatoshi Okutomi, Tokyo Institute of Technology; Kenta Endo, Hamamatsu Photonics K. K.; Yukinobu Sugiyama, Hamamatsu Photonics K. K.; Hajime Nagahara, Osaka University
P-3A-31	Consensus-Driven Propagation in Massive Unlabeled Data for Face Recognition	Xiaohang Zhan*, The Chinese University of Hong Kong; Ziwei Liu, The Chinese University of Hong Kong; Junjie Yan, Sensetime Group Limited; Dahua Lin, The Chinese University of Hong Kong; Chen Change Loy, Chinese University of Hong Kong
P-3A-32	Recovering Accurate 3D Human Pose in The Wild Using IMUs and a Moving Camera	Timo von Marcard*, University of Hannover; Roberto Henschel, Leibniz University of Hannover; Michael Black, Max Planck Institute for Intelligent Systems; Bodo Rosenhahn, Leibniz University Hannover; Gerard Pons-Moll, MPII, Germany
P-3A-33	Predicting Future Instance Segmentations by Forecasting Convolutional Features	Pauline Luc*, Facebook AI Research; Camille Couprie, Facebook; yann lecun, Facebook; Jakob Verbeek, INRIA
P-3A-34	PS-FCN: A Flexible Learning Framework for Photometric Stereo	Guanying Chen*, The University of Hong Kong; Kai Han, University of Oxford; Kwan-Yee Wong, The University of Hong Kong
P-3A-35	Unsupervised Class-Specific Deblurring	Nimisha T M*, Indian Institute of Technology Madras; Sunil Kumar, Indian Institute of Technology Madras; Rajagopalan Ambasamudram, Indian Institute of Technology Madras
P-3A-36	Face Super-resolution Guided by Facial Component Heatmaps	Xin Yu*, Australian National University; Basura Fernando, Australian National University; Bernard Ghanem, KAUST; Fatih Porikli, ANU; RICHARD HARTLEY, Australian National University, Australia

P-3A-37	A Contrario Horizon-First Vanishing Point Detection Using Second-Order Grouping Laws	Gilles Simon*, Université de Lorraine; Antoine Fond, Université de Lorraine; Marie-Odile Berger, INRIA
P-3A-38	Fast, Accurate, and, Lightweight Super-Resolution with Cascading Residual Network	Namhyuk Ahn, Ajou University; Byungkon Kang, Ajou University; Kyung-Ah Sohn*, Ajou University
P-3A-39	Face Recognition with Contrastive Convolution	Chunrui Han*, ICT, Chinese Academy of Sciences, China; Shiguang Shan, Chinese Academy of Sciences; Meina Kan, ICT, CAS; Shuzhe Wu, Chinese Academy of Sciences; xilin chen, ICT, Chinese Academy of Sciences, China
P-3A-40	Deforming Autoencoders: Unsupervised Disentangling of Shape and Appearance	Zhixin Shu*, Stony Brook University; Mihir Sahasrabudhe, CentraleSupelec; Alp Guler, INRIA; Dimitris Samaras, Stony Brook University; Nikos Paragios, Therapanacea; Iasonas Kokkinos , UCL
P-3A-41	NetAdapt: Platform-Aware Neural Network Adaptation for Mobile Applications	Tien-Ju Yang*, Massachusetts Institute of Technology; Andrew Howard, Google; Bo Chen, Google; Xiao Zhang, Google; Alec Go, Google; Vivienne Sze, Massachusetts Institute of Technology; Hartwig Adam, Google
P-3A-42	ExFuse: Enhancing Feature Fusion for Semantic Segmentation	Zhenli Zhang*, Fudan University; Xiangyu Zhang, Megvii Inc; Chao Peng, Megvii(Face++) Inc; Jian Sun, Megvii, Face++
P-3A-43	AugGAN: Cross Domain Adaptation with GAN-based Data Augmentation	Sheng-Wei Huang, National Tsing Hua University; Che-Tsung Lin*, National Tsing Hua University; Shu-Ping Chen, National Tsing Hua University; Yen-Yi Wu, NTHU CS; Po-Hao Hsu, National Tsing Hua University; Shang-Hong Lai , National Tsing Hua University

P-3A-44	LAPCSR:A Deep Laplacian Pyramid Generative Adversarial Network for Scalable Compressive Sensing Reconstruction	Kai Xu*, Arizona State University; Zhikang Zhang, Arizona State University; Fengbo Ren, Arizona State University
P-3A-45	U-PC: Unsupervised Planogram Compliance	Archan Ray, University of Massachusetts Amherst; Nishant Kumar, SMART-FM; Avishek Shaw*, Tata Consultancy Services Limited; Dipti Prasad Mukherjee, ISI, Kolkata
P-3A-46	Seeing Tree Structure from Vibration	Tianfan Xue, MIT; Jiajun Wu*, MIT; Zhoutong Zhang, MIT; Chengkai Zhang, MIT; Joshua Tenenbaum, MIT; Bill Freeman, MIT
P-3A-47	A Dataset of Flash and Ambient Illumination Pairs from the Crowd	Yagiz Aksoy*, ETH Zurich; Changil Kim, MIT CSAIL; Petr Kellnhofer, MIT; Sylvain Paris, Adobe Research; Mohamed A. Elghareb, Qatar Computing Research Institute; Marc Pollefeys, ETH Zurich; Wojciech Matusik, MIT
P-3A-48	Compressing the Input for CNNs with the First-Order Scattering Transform	Edouard Oyallon*, CentraleSupélec; Eugene Belilovsky, Inria Galen / KU Leuven; Sergey Zagoruyko, Inria; Michal Valko, Inria
P-3A-49	Distractor-aware Siamese Networks for Visual Object Tracking	Zheng Zhu*, CASIA; Qiang Wang, University of Chinese Academy of Sciences; Bo Li, sensetime; Wu Wei, Sensetime; Junjie Yan, Sensetime Group Limited
P-3A-50	„Factual“ or „Emotional“: Stylized Image Captioning with Adaptive Learning and Attention“	Tianlang Chen*, University of Rochester; Zhongping Zhang, University of Rochester; Quanzeng You, Microsoft; CHEN FANG, Adobe Research, San Jose, CA; Zhaowen Wang, Adobe Research; Hailin Jin, Adobe Research; Jiebo Luo, U. Rochester
P-3A-51	Constrained Optimization Based Low-Rank Approximation of Deep Neural Networks	Chong Li*, University of Washington; C.J. Richard Shi, University of Washington

P-3A-52	Extending Layered Models to 3D Motion	Dong Lao, KAUST; Ganesh Sundaramoorthi*, Kaust
P-3A-53	ExplainGAN: Model Explanation via Decision Boundary Crossing Transformations	Nathan Silberman*, Butterfly Network; Pouya Samangouei, Butterfly Network; Liam Nakagawa, Butterfly Network; Ardavan Saeedi, Butterfly Network Inc
P-3A-54	Adding Attentiveness to the Neurons in Recurrent Neural Networks	Pengfei Zhang, Xi'an Jiaotong University; Jianru Xue, Xi'an Jiaotong University; Cuiling Lan*, Microsoft Research; Wenjun Zeng, Microsoft Research; Zhanning Gao, Xi'an Jiaotong University; Nanning Zheng, Xi'an Jiaotong University
P-3A-55	ESPNet: Efficient Spatial Pyramid of Dilated Convolutions for Semantic Segmentation	Sachin Mehta*, University of Washington; Mohammad Rastegari, Allen Institute for Artificial Intelligence; Anat Caspi, University of Washington; Linda Shapiro, University of Washington; Hannaneh Hajishirzi, University of Washington
P-3A-56	Learning Human-Object Interactions by Graph Parsing Neural Networks	Siyuan Qi*, UCLA; Wenguan Wang, Beijing Institute of Technology; Baoxiong Jia, UCLA; Jianbing Shen, Beijing Institute of Technology; Song-Chun Zhu, UCLA

P-3A-57	PESTO: 6D Object Pose Estimation Benchmark	Tomas Hodan*, Czech Technical University in Prague; Frank Michel, Technical University Dresden; Eric Brachmann, TU Dresden; Wadim Kehl, Toyota Research Institute; Anders Buch, University of Southern Denmark; Dirk Kraft, Syddansk Universitet; Bertram Drost, MVTec Software GmbH; Joel Vidal, National Taiwan University of Science and Technology; Stephan Ihrke , Fraunhofer ivi ; Xenophon Zabulis, FORTH; Caner Sahin, Imperial College London; Fabian Manhardt, TU Munich; Federico Tombari, Technical University of Munich, Germany; Tae-Kyun Kim, Imperial College London; Jiri Matas, CMP CTU FEE; Carsten Rother, University of Heidelberg
P-3A-58	RCAA: Relational Context-Aware Agents for Person Search	Xiaojun Chang*, Carnegie Mellon University; Po-Yao Huang, Carnegie Mellon University; Xiaodan Liang, Carnegie Mellon University; Yi Yang, UTS; Alexander Hauptmann, Carnegie Mellon University
P-3A-59	DetNet: Design Backbone for Object Detection	Zeming Li*, Tsinghua University;Megvii Inc; Chao Peng, Megvii(-Face++) Inc; Gang Yu, Face++; Yangdong Deng, Tsinghua University; Xiangyu Zhang, Megvii Inc; Jian Sun, Megvii, Face++
P-3A-60	Modeling Varying Camera-IMU Time Offset in Optimization-Based Visual-Inertial Odometry	Yonggen Ling*, Tencent AI Lab; Linchao Bao, Tencent AI Lab; Zequn Jie, Tencent AI Lab; Fengming Zhu, Tencent AI Lab; Ziyang Li, Tencent AI Lab; Shanmin Tang, Tencent AI Lab; Yongsheng Liu, Tencent AI Lab; Wei Liu, Tencent AI Lab; Tong Zhang, Tecent AI Lab
P-3A-61	Exploiting temporal information for 3D human pose estimation	Mir Rayat Imtiaz Hossain*, University of British Columbia; Jim Little, University of British Columbia, Canada

P-3A-62	Joint Representation and Truncated Inference Learning for Correlation Filter based Tracking	Yingjie Yao, Harbin Institute of technology; Xiaohe Wu, Harbin Institute of technology; Lei Zhang, University of Pittsburgh; Shiguang Shan, Chinese Academy of Sciences; Wangmeng Zuo*, Harbin Institute of Technology, China
P-3A-63	Learning to Zoom: a Saliency-Based Sampling Layer for Neural Networks	Adria Recasens*, Massachusetts Institute of Technology; Petr Kellnhofer, MIT; Simon Stent, Toyota Research Institute; Wojciech Matusik, MIT; Antonio Torralba, MIT
P-3A-64	Does Haze Removal Help Image Classification?	Yanting Pei*, Beijing Jiaotong University; Yaping Huang, Beijing Jiaotong University; Qi Zou, Beijing Jiaotong University; Yuhang Lu, University of South Carolina; Song Wang, University of South Carolina
P-3A-65	Learning Local Descriptors by Integrating Geometry Constraints	Zixin Luo*, HKUST; Tianwei Shen, HKUST; Lei Zhou, HKUST; Siyu Zhu, HKUST; Runze Zhang, HKUST; Tian Fang, HKUST; Long Quan, Hong Kong University of Science and Technology
P-3A-66	Repeatability Is Not Enough: Learning Affine Regions via Discriminability	Dmytro Mishkin*, Czech Technical University in Prague; Filip Radenovic, Visual Recognition Group, CTU Prague; Jiri Matas, CMP CTU FEE
P-3A-67	Macro-Micro Adversarial Network for Human Parsing	Yawei Luo*, University of Technology Sydney; Zhedong Zheng, University of Technology Sydney; Liang Zheng, University of Technology Sydney; Yi Yang, UTS
P-3A-68	Learning Class Prototypes via Structure Alignment for Zero-Shot Recognition	Huajie Jiang, ICT, CAS; Ruiping Wang*, ICT, CAS; Shiguang Shan, Chinese Academy of Sciences; Xilin Chen, China
P-3A-69	SphereNet: Learning Spherical Representations for Detection and Classification in Omnidirectional Images	Benjamin Coors*, MPI Intelligent Systems, Bosch; Alexandru Condurache, Bosch; Andreas Geiger, MPI-IS and University of Tuebingen

P-3A-70	A dataset and architecture for visual reasoning with a working memory	Guangyu Robert Yang*, Columbia University; Igor Ganichev, Google Brain; Xiao-Jing Wang, New York University; Jon Shlens, Google; David Sussillo, Google Brain
P-3A-71	Flow-Grounded Spatial-Temporal Video Prediction from Still Images	Yijun Li*, University of California, Merced; CHEN FANG, Adobe Research, San Jose, CA; Jimei Yang, Adobe; Zhaowen Wang, Adobe Research; Xin Lu, Adobe; Ming-Hsuan Yang, University of California at Merced
P-3A-72	The Unmanned Aerial Vehicle Benchmark: Object Detection and Tracking	Dawei Du*, University of Chinese Academy of Sciences; Yuankai Qi, Harbin Institute of Technology; Hongyang Yu, Harbin Institute of Technology; Yifang Yang, University of Chinese Academy of Sciences; Kaiwen Duan, University of Chinese Academy of Sciences; guorong Li, CAS; Weigang Zhang, Harbin Institute of Technology, Weihai; Qingming Huang, University of Chinese Academy of Sciences; Qi Tian , The University of Texas at San Antonio
P-3A-73	Selective Zero-Shot Classification with Augmented Attributes	Jie Song, College of Computer Science and Technology, Zhejiang University; Chengchao Shen, Zhejiang University; Jie Lei, Zhejiang University; An-Xiang Zeng, Alibaba; Kairi Ou, Alibaba; Dacheng Tao, University of Sydney; Mingli Song*, Zhejiang University
P-3A-74	Action Search: Spotting Actions in Videos and Its Application to Temporal Action Localization	Humam Alwassel*, KAUST; Fabian Caba, KAUST; Bernard Ghanem, KAUST
P-3A-75	A Principled Approach to Hard Triplet Generation via Adversarial Nets	Yiru Zhao*, Shanghai Jiao Tong University; Zhongming Jin, Alibaba Group; Guo-Jun Qi, University of Central Florida; Hongtao Lu, Shanghai Jiao Tong University; Xian-Sheng Hua, Alibaba Group

P-3A-76	Pose Guided Human Video Generation	Ceyuan Yang*, SenseTime Group Limited; Zhe Wang, Sensetime Group Limited; Xinge Zhu, Sensetime Group Limited; Chen Huang, Carnegie Mellon University; Jianping Shi, Sensetime Group Limited; Dahua Lin, The Chinese University of Hong Kong
P-3A-77	Deep Directional Statistics: Pose Estimation with Uncertainty Quantification	Sergey Prokudin*, Max Planck Institute for Intelligent Systems; Sebastian Nowozin, Microsoft Research Cambridge; Peter Gehler, Amazon
P-3A-78	Learning 3D Human Pose from Structure and Motion	Rishabh Dabral*, IIT Bombay; Anurag Mundhada, IIT Bombay; Abhishek Sharma, Gobasco AI Labs
P-3A-79	Learning Dynamic Memory Networks for Object Tracking	Tianyu Yang*, City University of Hong Kong; Antoni Chan, City University of Hong Kong, Hong, Kong
P-3A-80	Faces as Lighting Probes via Unsupervised Deep Highlight Extraction	Renjiao Yi*, Simon Fraser University; Chenyang Zhu, Simon Fraser University; Ping Tan, Simon Fraser University; Stephen Lin, Microsoft Research
P-3A-81	CurriculumNet: Learning from Large-Scale Web Images without Human Annotation	Sheng Guo*, Malong Technologies; Weilin Huang, Malong Technologies; Haozhi Zhang, Malong Technologies
P-3A-82	Joint Task-Recursive Learning for Semantic Segmentation and Depth Estimation	Zhenyu Zhang*, Nanjing University of Sci & Tech; Zhen Cui, Nanjing University of Science and Technology; Zequn Jie, Tencent AI Lab; Xiang Li, NJUST; Chunyan Xu, Nanjing University of Science and Technology; Jian Yang, Nanjing University of Science and Technology
P-3A-83	HybridFusion: Real-Time Performance Capture Using a Single Depth Sensor and Sparse IMUs	Zerong Zheng*, Tsinghua University; Tao Yu, Beihang University; Hao Li, Pinscreen/University of Southern California/USC ICT; Kaiwen Guo, Google Inc.; Qionghai Dai, Tsinghua University; Lu Fang, Tsinghua University; Yebin Liu, Tsinghua University

P-3A-84	Associating Inter-Image Salient Instances for Weakly Supervised Semantic Segmentation	Ruochen Fan*, Tsinghua University; Qibin Hou, Nankai University; Ming-Ming Cheng, Nankai University; Gang Yu, Face++; Ralph Martin, Cardiff University; Shimin Hu, Tsinghua University
P-3A-85	Ask, Acquire and Attack: Data-free UAP generation using Class impressions	Konda Reddy Mopuri*, Indian Institute of Science, Bangalore; Phani Krishna Uppala, Indian Institute of Science; Venkatesh Babu RADHAKRISHNAN, Indian Institute of Science
P-3A-86	A Scalable Exemplar-based Subspace Clustering Algorithm for Class-Imbalanced Data	Chong You*, Johns Hopkins University; Chi Li, Johns Hopkins University; Daniel Robinson, Johns Hopkins University; Rene Vidal, Johns Hopkins University
P-3A-87	Find and Focus: Retrieve and Localize Video Events with Natural Language Queries	Dian SHAO*, The Chinese University of Hong Kong; Yu Xiong, The Chinese University of HK; Yue Zhao, The Chinese University of Hong Kong; Qingqiu Huang, CUHK; Yu Qiao, Multimedia Laboratory, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences; Dahua Lin, The Chinese University of Hong Kong
P-3A-88	Graininess-Aware Deep Feature Learning for Pedestrian Detection	Chunze Lin, Tsinghua University; Jiwen Lu*, Tsinghua University; Jie Zhou, Tsinghua University, China
P-3A-89	MVSNet: Depth Inference for Unstructured Multi-view Stereo	Yao Yao*, The Hong Kong University of Science and Technology; Zixin Luo, HKUST; Shiwei Li, HKUST; Tian Fang, HKUST; Long Quan, Hong Kong University of Science and Technology
P-3A-90	PlaneMatch: Patch Coplanarity Prediction for Robust RGB-D Registration	Yifei Shi, Princeton University; Kai Xu, Princeton University and National University of Defense Technology; Matthias Niessner, Technical University of Munich; Szymon Rusinkiewicz, Princeton University; Thomas Funkhouser*, Princeton, USA

P-3A-91	Deep Virtual Stereo Odometry: Leveraging Deep Depth Prediction for Monocular Direct Sparse Odometry	Nan Yang*, Technical University of Munich; Rui Wang, Technical University of Munich; Joerg Stueckler, Technical University of Munich; Daniel Cremers, TUM
P-3B-01	GANimation: Anatomically-aware Facial Animation from a Single Image	Albert Pumarola*, Institut de Robotica i Informatica Industrial; Antonio Agudo, Institut de Robotica i Informatica Industrial, CSIC-UPC; Aleix Martinez, The Ohio State University; Alberto Sanfeliu, Industrial Robotics Institute; Francesc Moreno, IRI
P-3B-02	Unsupervised Geometry-Aware Representation for 3D Human Pose Estimation	Helge Rhodin*, EPFL; Mathieu Salzmann, EPFL; Pascal Fua, EPFL, Switzerland
P-3B-03	Efficient Semantic Scene Completion Network with Spatial Group Convolution	Jiahui Zhang*, Tsinghua University; Hao Zhao, Intel Labs China; Anbang Yao, Intel Labs China; Yurong Chen, Intel Labs China; Hongen Liao, Tsinghua University
P-3B-04	Deep Autoencoder for Combined Human Pose Estimation and Body Model Upscaling	Matthew Trumble*, University of Surrey; Andrew Gilbert, University of Surrey; John Collomosse, Adobe Research; Adrian Hilton, University of Surrey
P-3B-05	Highly-Economized Multi-View Binary Compression for Scalable Image Clustering	Zheng Zhang*, Harbin Institute of Technology Shenzhen Graduate School; Li Liu, the inception institute of artificial intelligence; Jie Qin, ETH Zurich; Fan Zhu, the inception institute of artificial intelligence ; Fumin Shen, UESTC; Yong Xu, Harbin Institute of Technology Shenzhen Graduate School; Ling Shao, Inception Institute of Artificial Intelligence; Heng Tao Shen, University of Electronic Science and Technology of China (UESTC)

P-3B-06	Asynchronous, Photometric Feature Tracking using Events and Frames	Daniel Gehrig, University of Zurich; Henri Rebecq*, University of Zurich; Guillermo Gallego, University of Zurich; Davide Scaramuzza, University of Zurich& ETH Zurich, Switzerland
P-3B-07	Deterministic Consensus Maximization with Biconvex Programming	Zhipeng Cai*, The University of Adelaide; Tat-Jun Chin, University of Adelaide; Huu Le, University of Adelaide; David Suter, University of Adelaide
P-3B-08	Depth-aware CNN for RGB-D Segmentation	Weiyue Wang*, USC; Ulrich Neumann, USC
P-3B-09	Object Detection in Video with Spatiotemporal Sampling Networks	Gedas Bertasius*, University of Pennsylvania; Lorenzo Torresani, Dartmouth College; Jianbo Shi, University of Pennsylvania
P-3B-10	Dependency-aware Attention Control for Unconstrained Face Recognition with Image Sets	Xiaofeng Liu*, Carnegie Mellon University; B. V. K. Vijaya Kumar, CMU, USA; Chao Yang, University of Southern California; Qingming Tang, TTIC; Jane You, The Hong Kong Polytechnic University
P-3B-11	License Plate Detection and Recognition in Unconstrained Scenarios	Sérgio Silva*, UFRGS; Claudio Jung, UFRGS
P-3B-12	Revisiting the Inverted Indices for Billion-Scale Approximate Nearest Neighbors	Dmitry Baranchuk*, MSU / Yandex; Artem Babenko, MIPT/Yandex; Yury Malkov, NTechLab
P-3B-13	Zero-Annotation Object Detection with Web Knowledge Transfer	Qingyi Tao*, Nanyang Techonological University; Hao Yang, NTU; Jianfei Cai, Nanyang Technological University
P-3B-14	Semi-supervised Adversarial Learning to Generate Photorealistic Face Images of New Identities from 3D Morphable Model	Baris Gecer*, Imperial College London; Binod Bhattarai, Imperial College London; Josef Kittler, University of Surrey, UK; Tae-Kyun Kim, Imperial College London

P-3B-15	Improving Shape Deformation in Unsupervised Image-to-Image Translation	Aaron Gokaslan*, Brown University; Vivek Ramanujan, Brown University; Daniel Ritchie, Brown University; Kwang In Kim, University of Bath; James Tompkin, Brown University
P-3B-16	K-convexity shape priors for segmentation	Hossam Isack*, UWO; Lena Gorelick, University of Western Ontario; Karin nG, University of Western Ontario; Olga Veksler, University of Western Ontario; Yuri Boykov, University of Waterloo
P-3B-17	Visual Question Generation for Class Acquisition of Unknown Objects	Kohei Uehara*, The University of Tokyo; Antonio Tejero-de-Pablos, The University of Tokyo; Yoshitaka Ushiku, The University of Tokyo; Tatsuya Harada, The University of Tokyo
P-3B-18	Sampling Algebraic Varieties for Robust Camera Autocalibration	Danda Pani Paudel*, ETH Zürich; Luc Van Gool, ETH Zurich
P-3B-19	Hand Pose Estimation via Latent 2.5D Heatmap Regression	Umar Iqbal*, University of Bonn; Pavlo Molchanov, NVIDIA; Thomas Breuel, NVIDIA; Jürgen Gall, University of Bonn; Kautz Jan, NVIDIA
P-3B-20	HairNet: Single-View Hair Reconstruction using Convolutional Neural Networks	Yi Zhou*, University of Southern California; Liwen Hu, University of Southern California; Jun Xing, Institute for Creative Technologies, USC; Weikai Chen, USC Institute for Creative Technology; Han-Wei Kung, University of California, Santa Barbara; Xin Tong, Microsoft Research Asia; Hao Li, Pinscreen/University of Southern California/USC ICT
P-3B-21	Super-Identity Convolutional Neural Network for Face Hallucination	Kaipeng Zhang*, National Taiwan University; ZHANPENG ZHANG, SenseTime Group Limited; Chia-Wen Cheng, UT Austin; Winston Hsu, National Taiwan University; Yu Qiao, Multimedia Laboratory, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences; Wei Liu, Tencent AI Lab; Tong Zhang, Tecent AI Lab

P-3B-22	Receptive Field Block Net for Accurate and Fast Object Detection	Songtao Liu, BUAA; Di Huang*, Beihang University, China; Yunhong Wang, State Key Laboratory of Virtual Reality Technology and System, Beihang University, Beijing 100191, China
P-3B-23	Interpretable Intuitive Physics Model	Tian Ye*, Carnegie Mellon University; Xiaolong Wang, CMU; James Davidson, Google; Abhinav Gupta, CMU
P-3B-24	Variable Ring Light Imaging: Capturing Transient Subsurface Scattering with An Ordinary Camera	Ko Nishino*, Kyoto University; Art Subpa-asa, Tokyo Institute of Technology; Yuta Asano, Tokyo Institute of Technology; Mihoko Shimano, National Institute of Informatics; Imari Sato, National Institute of Informatics
P-3B-25	Facial Dynamics Interpreter Network: What are the Important Relations between Local Dynamics for Facial Trait Estimation?	Seong Tae Kim*, KAIST; Yong Man Ro, KAIST
P-3B-26	Coloring with Words: Guiding Image Colorization Through Text-based Palette Generation	Hyojin Bahng, Korea University; Seungjoo Yoo, Korea University; Wonwoong Cho, Korea University; David Park, Korea University; Ziming Wu, Hong Kong University of Science and Technology; Xiaojuan MA, Hong Kong University of Science and Technology; Jaegul Choo*, Korea University
P-3B-27	Sparsely Aggregated Convolutional Networks	Ligeng Zhu*, Simon Fraser University; Ruizhi Deng, Simon Fraser University; Michael Maire, Toyota Technological Institute at Chicago; Zhiwei Deng, Simon Fraser University; Greg Mori, Simon Fraser University; Ping Tan, Simon Fraser University
P-3B-28	Deep Attention Neural Tensor Network for Visual Question Answering	Yalong Bai*, Harbin Institute of Technology; Jianlong Fu, Microsoft Research; Tao Mei, JD.com

P-3B-29	Diverse feature visualizations reveal invariances in early layers of deep neural networks	Santiago Cadena*, University of Tübingen; Marissa Weis, University of Tübingen; Leon A. Gatys, University of Tuebingen; Matthias Bethge, University of Tübingen; Alexander Ecker, University of Tübingen
P-3B-30	Sidekick Policy Learning for Active Visual Exploration	Santhosh Kumar Ramakrishnan*, University of Texas at Austin; Kristen Grauman, University of Texas
P-3B-31	DPP-Net: Device-aware Progressive Search for Pareto-optimal Neural Architectures	Jin-Dong Dong*, National Tsing-Hua University; An-Chieh Cheng, National Tsing-Hua University; Da-Cheng Juan, Google; Wei Wei, Google; Min Sun, NTHU
P-3B-32	Pixel2Mesh: Generating 3D Mesh Models from Single RGB Images	Nanyang Wang, Fudan University; Yinda Zhang*, Princeton University; Zhuwen Li, Intel Labs; Yanwei Fu, Fudan Univ.; Wei Liu, Tencent AI Lab; Yu-Gang Jiang, Fudan University
P-3B-33	End-to-End Incremental Learning	Francisco M. Castro*, University of Málaga; Manuel J. Marín-Jiménez, University of Córdoba; Nicolás Guil, University of Málaga; Cordelia Schmid, INRIA; Karteek Alahari, Inria
P-3B-34	CAR-Net: Clairvoyant Attentive Recurrent Network	Amir Sadeghian*, Stanford; Maxime Voisin, Stanford University; Ferdinand Legros, Stanford University; Ricky Vesel, Race Optimal; Alexandre Alahi, EPFL; Silvio Savarese, Stanford University
P-3B-35	Learning Data Terms for Image Deblurring	Jiangxin Dong*, Dalian University of Technology; Jinshan Pan, Dalian University of Technology; Deqing Sun, NVIDIA; Zhixun Su, Dalian University of Technology; Ming-Hsuan Yang, University of California at Merced
P-3B-36	Image Inpainting for Irregular Holes Using Partial Convolutions	Guilin Liu*, NVIDIA; Fitsum Reda, NVIDIA; Kevin Shih, NVIDIA; Ting-Chun Wang, NVIDIA; Andrew Tao, NVIDIA; Bryan Catanzaro, NVIDIA

P-3B-37	SRDA: Generating Instance Segmentation Annotation Via Scanning, Reasoning And Domain Adaption	Wenqiang Xu, Shanghai Jiaotong University; Yonglu Li, Shanghai Jiao Tong University; Jun Lv, SJTU; Cewu Lu*, Shanghai Jiao Tong Univercity
P-3B-38	Learning Priors for Semantic 3D Reconstruction	Ian Cherabier*, ETH Zurich; Johannes Schoenberger, ETH Zurich; Martin R. Oswald, ETH Zurich; Marc Pollefeys, ETH Zurich; Andreas Geiger, MPI-IS and University of Tuebingen
P-3B-39	Integrating Egocentric Videos in Top-view Surveillance Videos: Joint Identification and Temporal Alignment	Shervin Ardeshir*, University of Central Florida; Ali Borji, University of Central Florida
P-3B-40	Deep Boosting for Image Denoising	Chang Chen, University of Science and Technology of China; Zhiwei Xiong*, University of Science and Technology of China; Xinmei Tian, USTC; Feng Wu, University of Science and Technology of China
P-3B-41	Descending, lifting or smoothing: Secrets of robust cost optimization	Christopher Zach*, Toshiba Research; Guillaume Bourmaud, University of Bordeaux
P-3B-42	MultiPoseNet: Fast Multi-Person Pose Estimation using Pose Residual Network	Muhammed Kocabas*, Middle East Technical University; Salih Karagoz, Middle East Technical University; Emre Akbas, Middle East Technical University
P-3B-43	TS2C: Tight Box Mining with Surrounding Segmentation Context for Weakly Supervised Object Detection	Yunchao Wei*, UIUC; Zhiqiang Shen, UIUC; Honghui Shi, UIUC; Bowen Cheng, UIUC; Jinjun Xiong, IBM Thomas J. Watson Research Center; Jiashi Feng, NUS; Thomas Huang, UIUC
P-3B-44	End-to-End Deep Structured Models for Drawing Crosswalks	Justin Liang*, Uber ATG; Raquel Urtasun, Uber ATG

P-3B-45	Efficient Global Point Cloud Registration by Matching Rotation Invariant Features Through Translation Search	Yinlong Liu, Fudan University; Wang Chen*, Shanghai Key Laboratory of Medical Imaging Computing and Computer Assisted Intervention, Digital Medical Research Center, Fudan University; Zhijian Song, Fudan University; Manning Wang, Fudan University
P-3B-46	Large Scale Urban Scene Modeling from MVS Meshes	Lingjie Zhu, University of Chinese Academy of Sciences; National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences; Shuhan Shen*, National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences; Zhanyi Hu, National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences
P-3B-47	Sub-GAN: An Unsupervised Generative Model via Subspaces	Jie Liang, Nankai University; Jufeng Yang*, Nankai University ; Hsin-Ying Lee, University of California, Merced; Kai Wang, Nankai University; Ming-Hsuan Yang, University of California at Merced
P-3B-48	Pseudo Pyramid Deeper Bidirectional ConvLSTM for Video Saliency Detection	Hongmei Song, Beijing Institute of Technology; Sanyuan Zhao*, Beijing Institute of Technology ; Jianbing Shen, Beijing Institute of Technology; Kin-Man Lam, The Hong Kong Polytechnic University
P-3B-49	Practical Black-box Attacks on Deep Neural Networks using Efficient Query Mechanisms	Arjun Nitin Bhagoji*, Princeton University; Warren he, University of California, Berkeley; Bo Li, University of Illinois at Urbana Champaign; Dawn Song, UC Berkeley
P-3B-50	Learning 3D Shape Priors for Shape Completion and Reconstruction	Jiajun Wu*, MIT; Chengkai Zhang, MIT; Xiuming Zhang, MIT; Zhoutong Zhang, MIT; Joshua Tenenbaum, MIT; Bill Freeman, MIT

P-3B-51	Comparator Networks	Weidi Xie*, University of Oxford; Li Shen, University of Oxford; Andrew Zisserman, University of Oxford
P-3B-52	Improving Fine-Grained Visual Classification using Pairwise Confusion	Abhimanyu Dubey*, Massachusetts Institute of Technology; Otkrist Gupta, MIT; Pei Guo, Brigham Young University; Ryan Farrell, Brigham Young University; Ramesh Raskar, Massachusetts Institute of Technology; Nikhil Naik, MIT
P-3B-53	Visual-Inertial Object Detection and Mapping	Xiaohan Fei*, UCLA; Stefano Soatto, UCLA
P-3B-54	Learning Region Features for Object Detection	Jiayuan Gu, Peking University; Han Hu, Microsoft Research Asia; Liwei Wang, Peking University; Yichen Wei, MSR Asia; Jifeng Dai*, Microsoft Research Asia
P-3B-55	Efficient Dense Point Cloud Object Reconstruction using Deformation Vector Fields	Kejie Li*, University of Adelaide; Trung Pham, NVIDIA; Huangying Zhan, The University of Adelaide; Ian Reid, University of Adelaide, Australia
P-3B-56	Evaluating Capability of Deep Neural Networks for Image Classification via Information Plane	Hao Cheng*, Shanghaitech University; Dongze Lian, Shanghaitech University; Shenghua Gao, Shanghaitech University; Yanlin Geng, Shanghaitech University
P-3B-57	Shuffle-Then-Assemble: Learning Object-Agnostic Visual Relationship Features	XU YANG*, NTU; Hanwang Zhang, Nanyang Technological University; Jianfei Cai, Nanyang Technological University
P-3B-58	Zero-Shot Deep Domain Adaptation	Kuan-Chuan Peng*, siemens corporation; Ziyan Wu, Siemens Corporation; Jan Ernst, Siemens Corporation
P-3B-59	Deep Imbalanced Attribute Classification using Visual Attention Aggregation	Nikolaos Sarafianos*, University of Houston; Xiang Xu, University of Houston; Ioannis Kakadiaris, University of Houston

P-3B-60	Video Object Segmentation by Learning Location-Sensitive Embeddings	Hai Ci, Peking University; Chunyu Wang*, Microsoft Research asia; Yizhou Wang, PKU
P-3B-61	Deep Multi-Task Learning to Recognise Subtle Facial Expressions of Mental States	Guosheng Hu*, AnyVision; Li Liu, the inception institute of artificial intelligence; Yang Yuan, AnyVision; Zehao Yu, Xiamen University; Yang Hua, Queen's University Belfast; Zhihong Zhang, Xiamen University; Fumin Shen, UESTC; Ling Shao, Inception Institute of Artificial Intelligence; Timothy Hospedales, Edinburgh University; Neil Robertson, Queen's University Belfast; Yongxin Yang, University of Edinburgh
P-3B-62	Where Will They Go? Predicting Fine-Grained Adversarial Multi-Agent Motion using Conditional Variational Autoencoders	Panna Felsen*, University of California Berkeley; Patrick Lucey, STATS; Sujoy Ganguly, STATS
P-3B-63	Video Summarization Using Fully Convolutional Sequence Networks	Mrigank Rochan*, University of Manitoba; Linwei Ye, University of Manitoba; Yang Wang, University of Manitoba
P-3B-64	Occlusion-aware Hand Pose Estimation Using Hierarchical Mixture Density Network	Qi Ye*, Imperial College London; Tae-Kyun Kim, Imperial College London
P-3B-65	Learning with Biased Complementary Labels	Xiyu Yu*, The University of Sydney; Tongliang Liu, The University of Sydney; Mingming Gong, University of Pittsburgh; Dacheng Tao, University of Sydney
P-3B-66	ConceptMask: Large-Scale Segmentation from Semantic Concepts	Yufei Wang*, Facebook; Zhe Lin, Adobe Research; Xiaohui Shen, Adobe Research; Scott Cohen, Adobe Research; Jianming Zhang, Adobe Research

P-3B-67	Conditional Image-Text Embedding Networks	Bryan Plummer*, Boston University; Paige Kordas, University of Illinois at Urbana Champaign; Hadi Kiapour, eBay; Shuai Zheng, eBay; Robinson Piramuthu, eBay Inc.; Svetlana Lazebnik, UIUC
P-3B-68	Geolocation Estimation of Photos using a Hierarchical Model and Scene Classification	Eric Müller-Budack*, Leibniz Information Centre of Science and Technology (TIB); Kader Pustu-Iren, Leibniz Information Center of Science and Technology (TIB); Ralph Ewerth, Leibniz Information Center of Science and Technology (TIB)
P-3B-69	Lifting Layers: Analysis and Applications	Michael Moeller*, University of Siegen; Peter Ochs, Saarland University; Tim Meinhardt, Technical University of Munich; Laura Leal-Taixé, TUM
P-3B-70	Progressive Neural Architecture Search	Chenxi Liu*, Johns Hopkins University; Maxim Neumann, Google; Barret Zoph, Google; Jon Shlens, Google; Wei Hua, Google; Li-Jia Li, Google; Li Fei-Fei, Stanford University; Alan Yuille, Johns Hopkins University; Jonathan Huang, Google; Kevin Murphy, Google
P-3B-71	Learning Deep Representations with Probabilistic Knowledge Transfer	Nikolaos Passalis*, Aristotle University of Thessaloniki; Anastasios Tefas, Aristotle University of Thessaloniki
P-3B-72	Robust fitting in computer vision: easy or hard?	Tat-Jun Chin*, University of Adelaide; Zhipeng Cai, The University of Adelaide; Frank Neumann, The University of Adelaide, School of Computer Science, Faculty of Engineering, Computer and Mathematical Science
P-3B-73	Dual-Agent Deep Reinforcement Learning for Deformable Face Tracking	Minghao Guo, Tsinghua University; Jiwen Lu*, Tsinghua University; Jie Zhou, Tsinghua University, China

P-3C-01	Zero-Shot Object Detection	Ankan Bansal*, University of Maryland; Karan Sikka, SRI International; Gaurav Sharma, NEC Labs America; Rama Chellappa, University of Maryland; Ajay Divakaran, SRI, USA
P-3C-02	ForestHash: Semantic Hashing With Shallow Random Forests and Tiny Convolutional Networks	Qiang Qiu*, Duke University; Jose Lezama, Universidad de la Republica, Uruguay; Alex Bronstein, Tel Aviv University, Israel; Guillermo Sapiro, Duke University
P-3C-03	ML-LocNet: Improving Object Localization with Multi-view Learning Network	Xiaopeng Zhang*, National University of Singapore; Jiashi Feng, NUS
P-3C-04	MPLP++: Fast, Parallel Dual Block-Coordinate Ascent for Dense Graphical Models	Siddharth Tourani*, Visual Learning Lab, HCI, Uni-Heidelberg; Alexander Shekhovtsov, Czech Technical University in Prague, Czech Republic; Carsten Rother, University of Heidelberg; Bogdan Savchynskyy, Heidelberg University
P-3C-05	A Zero-Shot Framework for Sketch based Image Retrieval	Sasikiran Yelamarthi , IIT Madras; Shiva Krishna Reddy M, Indian Institute of Technology Madras; Ashish Mishra*, IIT Madras; Anurag Mittal, Indian Institute of Technology Madras
P-3C-06	In the Eye of Beholder: Joint Learning of Gaze and Actions in First Person Vision	Yin Li*, CMU; Miao Liu, Georgia Tech; James Rehg, Georgia Institute of Technology
P-3C-07	SAN: Learning Relationship between Convolutional Features for Multi-Scale Object Detection	YongHyun Kim*, POSTECH
P-3C-08	A Systematic DNN Weight Pruning Framework using Alternating Direction Method of Multipliers	Tianyun Zhang*, Syracuse University; Shaokai Ye, Syracuse University; Kaiqi Zhang, Syracuse University; Yanzhi Wang, Syracuse University; Makan Fardad, Syracuse Universtiy; Wujie Wen, Florida International University

P-3C-09	Iterative Crowd Counting	Viresh Ranjan*, Stony Brook University; Hieu Le, Stony Brook University; Minh Hoai Nguyen, Stony Brook University
P-3C-10	A Dataset for Lane Instance Segmentation in Urban Environments	Brook Roberts, Five AI Ltd.; Sebastian Kaltwang*, Five AI Ltd.; Sina Samangooei, Five AI Ltd.; Mark Pender-Bare, Five AI Ltd.; Konstantinos Tertikas, Five AI Ltd.; John Redford, Five AI Ltd.
P-3C-11	Out-of-Distribution Detection Using an Ensemble of Self Supervised Leave-out Classifiers	Nataraj Jammalamadaka*, Intel Labs; Xia Zhu, Intel Labs; Dipankar Das, Intel Labs; Bharat Kaul, Intel Labs; Theodore Willke, Intel Labs
P-3C-12	Penalizing Top Performers: Conservative Loss for Semantic Segmentation Adaptation	Xinge Zhu*, Sensetime Group Limited; Hui Zhou, Sensetime Group Limited.; Ceyuan Yang, SenseTime Group Limited; Jianping Shi, Sensetime Group Limited; Dahua Lin, The Chinese University of Hong Kong
P-3C-13	Compound Memory Networks for Few-shot Video Classification	Linchao Zhu*, University of Technology, Sydney; Yi Yang, UTS
P-3C-14	Straight to the Facts: Learning Knowledge Base Retrieval for Factual Visual Question Answering	Medhini Narasimhan*, University of Illinois at Urbana-Champaign ; Alexander Schwing, UIUC
P-3C-15	Interpretable Basis Decomposition for Visual Explanation	Antonio Torralba, MIT; Bolei Zhou*, MIT; David Bau, MIT; Yiyou Sun, Harvard
P-3C-16	How Local is the Local Diversity? Reinforcing Sequential Determinantal Point Processes with Dynamic Ground Sets for Supervised Video Summarization	Yandong Li*, University of Central Florida; Boqing Gong, Tencent AI Lab; Tianbao Yang, University of Iowa; Liqiang Wang, University of Central Florida

P-3C-17	Dividing and Aggregating Network for Multi-view Action Recognition	Dongang Wang*, The University of Sydney; Wanli Ouyang, CUHK; Wen Li, ETHZ; Dong Xu, University of Sydney
P-3C-18	Shape Reconstruction Using Volume Sweeping and Learned Photoconsistency	Vincent Leroy*, INRIA Grenoble Rhône-Alpes; Edmond Boyer, Inria; Jean-Sebastien Franco, INRIA
P-3C-19	RT-GENE: Real-Time Eye Gaze Estimation in Natural Environments	Tobias Fischer*, Imperial College London; Hyung Jin Chang, University of Birmingham; Yiannis Demiris, Imperial College London
P-3C-20	Pairwise Body-Part Attention for Recognizing Human-Object Interactions	Haoshu Fang, SJTU; Jinkun Cao, Shanghai Jiao Tong University; Yu-Wing Tai, Tencent YouTu; Cewu Lu*, Shanghai Jiao Tong Univercity
P-3C-21	Motion Feature Network: Fixed Motion Filter for Action Recognition	Myunggi Lee, Seoul National University; Seung Eui Lee, Seoul National University; Sung Joon Son, Seoul National University; Gyutae Park, Seoul National University; Nojun Kwak*, Seoul National University
P-3C-22	Reverse Attention for Salient Object Detection	Shuhan Chen*, Yangzhou University; Xiuli Tan, Yangzhou University; Ben Wang, Yangzhou University; Xuelong Hu, Yangzhou University
P-3C-23	Dynamic Sampling Convolutional Neural Networks	Jialin Wu*, UT Austin; Dai Li, Tsinghua University; Yu Yang, Tsinghua University; Chandrajit Bajaj, University of Texas, Austin; Xiangyang Ji, Tsinghua University
P-3C-24	DDRNet: Depth Map Denoising and Refinement for Consumer Depth Cameras Using Cascaded CNNs	Shi Yan, Tsinghua University; Chenglei Wu, Oculus Research; Lizheng Wang, Tsinghua University; Liang An, Tsinghua University; Feng Xu, Tsinghua University; Kaiwen Guo, Google Inc.; Yebin Liu*, Tsinghua University

P-3C-25	Stereo Computation for a Single Mixture Image	Yiran Zhong, Australian National University; Yuchao Dai*, Northwestern Polytechnical University; HONGDONG LI, Australian National University, Australia
P-3C-26	Volumetric performance capture from minimal camera viewpoints	Andrew Gilbert*, University of Surrey; Marco Volino, University of Surrey; John Collomosse, Adobe Research; Adrian Hilton, University of Surrey
P-3C-27	Liquid Pouring Monitoring via Rich Sensory Inputs	Tz-Ying Wu*, National Tsing Hua University; Juan-Ting Lin, National Tsing Hua University; Tsun-Hsuang Wang, National Tsing Hua University; Chan-Wei Hu, National Tsing Hua University; Juan Carlos Niebles, Stanford University; Min Sun, NTHU
P-3C-28	Move Forward and Tell: A Progressive Generator of Video Descriptions	Yilei Xiong*, The Chinese University of Hong Kong; Bo Dai, the Chinese University of Hong Kong; Dahua Lin, The Chinese University of Hong Kong
P-3C-29	DYAN: A Dynamical Atoms-Based Network for Video Prediction	Wenqian Liu*, Northeastern University; Abhishek Sharma, Northeastern University ; Octavia Camps, Northeastern University; Mario Sznaier, Northeastern University
P-3C-30	Deep Structure Inference Network for Facial Action Unit Recognition	Ciprian Corneanu*, Universitat de Barcelona; Meysam Madadi, CVC; Sergio Escalera, Computer Vision Center (UAB) & University of Barcelona,
P-3C-31	Physical Primitive Decomposition	Zhijian Liu, Shanghai Jiao Tong University; Jiajun Wu*, MIT; Bill Freeman, MIT; Joshua Tenenbaum, MIT
P-3C-32	Boosted Attention: Leveraging Human Attention for Image Captioning	Shi Chen*, University of Minnesota; Qi Zhao, University of Minnesota
P-3C-33	Is Robustness the Cost of Accuracy? -- Lessons Learned from 18 Deep Image Classifiers	Dong Su*, IBM Research T.J. Watson Center; Huan Zhang, UC Davis; Hongge Chen, MIT; Jinfeng Yi, JD AI Research; Pin-Yu Chen, IBM Research; Yupeng Gao, IBM Research AI

P-3C-34	Dynamic Multimodal Instance Segmentation guided by natural language queries	Edgar Margffoy-Tuay*, Universidad de los Andes; Emilio Botero, Universidad de los Andes; Juan Pérez, Universidad de los Andes; PABLO ARBELÁEZ, Universidad de los Andes
P-3C-35	Hierarchy of Alternating Specialists for Scene Recognition	Hyo Jin Kim*, University of North Carolina at Chapel Hill; Jan-Michael Frahm, UNC-Chapel Hill
P-3C-36	SwapNet: Garment Transfer in Single View Images	Amit Raj*, Georgia Institute of Technology; Patsorn Sangkloy, Georgia Institute of Technology; Huiwen Chang, Princeton University; Jingwan Lu, Adobe Research ; Duygu Ceylan, Adobe Research; James Hays, Georgia Institute of Technology, USA
P-3C-37	What do I Annotate Next? An Empirical Study of Active Learning for Action Localization	Fabian Caba*, KAUST; Joon-Young Lee, Adobe Research; Hailin Jin, Adobe Research; Bernard Ghanem, KAUST
P-3C-38	Combining 3D Model Contour Energy and Keypoints for Object Tracking	Bogdan Bugaev*, Saint Petersburg Academic University; Anton Kryshchenko, Saint Petersburg Academic University; Roman Belov, KeenTools
P-3C-39	AGIL: Learning Attention from Human for Visuomotor Tasks	Ruohan Zhang*, University of Texas at Austin; Zhuode Liu, Google Inc.; Luxin Zhang, Peking University; Jake Whritner, University of Texas at Austin; Karl Muller, University of Texas at Austin; Mary Hayhoe, University of Texas at Austin; Dana Ballard, University of Texas at Austin
P-3C-40	PersonLab: Person Pose Estimation and Instance Segmentation with a Bottom-Up, Part-Based, Geometric Embedding Model	George Papandreou*, Google; Tyler Zhu, Google; Liang-Chieh Chen, Google Inc.; Spyros Gidaris, Ecole des Ponts ParisTech; Jonathan Tompson, Google; Kevin Murphy, Google

P-3C-41	Accelerating Dynamic Programs via Nested Benders Decomposition with Application to Multi-Person Pose Estimation	Shaofei Wang*, Baidu Inc.; Alexander Ihler, UC Irvine; Konrad Kording, Northwestern; Julian Yarkony, Experian Data Lab
P-3C-42	Separating Reflection and Transmission Images in the Wild	Patrick Wieschollek*, University of Tuebingen; Orazio Gallo, NVIDIA Research; Jinwei Gu, Nvidia; Kautz Jan, NVIDIA
P-3C-43	Point-to-Point Regression PointNet for 3D Hand Pose Estimation	Liuhao Ge*, NTU; Zhou Ren, Snap Research, USA, ; Junsong Yuan, State University of New York at Buffalo, USA
P-3C-44	Summarizing First-Person Videos from Third Persons' Points of View	HSUAN-I HO*, National Taiwan University; Wei-Chen Chiu, National Chiao Tung University; Yu-Chiang Frank Wang, National Taiwan University
P-3C-45	Learning Category-Specific Mesh Reconstruction from Image Collections	Angjoo Kanazawa*, UC Berkeley; Shubham Tulsiani, UC Berkeley; Alexei Efros, UC Berkeley; Jitendra Malik, University of California at Berkley
P-3C-46	StereoNet: Guided Hierarchical Refinement for Real-Time Edge-Aware Depth Prediction	Sameh Khamis*, Google; Sean Fanello, Google; Christoph Rhemann, Google; Julien Valentin, Google; Adarsh Kowdle, Google; Shahram Izadi, Google
P-3C-47	Visual Question Answering as a Meta Learning Task	Damien Teney*, The Unversity of Adelaide; Anton van den Hengel, The University of Adelaide
P-3C-48	SRFeat: Single Image Super Resolution with Feature Discrimination	Seong-Jin Park*, POSTECH; Hyeongseok Son, POSTECH; Sunghyun Cho, DGIST; Ki-Sang Hong, POSTECH; Seungyong Lee, POSTECH
P-3C-49	Deep Factorised Inverse-Sketching	Kaiyue Pang*, Queen Mary University of London; Da Li, QMUL; Jifei Song, Queen Mary, University of London; Yi-Zhe Song, Queen Mary University of London; Tao Xiang, Queen Mary, University of London, UK; Timothy Hospedales, Edinburgh University

P-3C-50	Multimodal image alignment through a multiscale chain of neural networks with application to remote sensing	Armand Zampieri, Inria Sophia-Antipolis; Guillaume Charpiat, INRIA; Nicolas Girard, Inria Sophia-Antipolis; Yuliya Tarabalka*, Inria Sophia-Antipolis
P-3C-51	Improving Deep Visual Representation for Person Re-identification by Global and Local Image-language Association	Dapeng Chen*, The Chinese University of HongKong; Hongsheng Li, Chinese University of Hong Kong; Xihui Liu, The Chinese University of Hong Kong; Jing Shao, The Chinese University of Hong Kong; Xiaogang Wang, Chinese University of Hong Kong, Hong Kong
P-3C-52	Robust Optical Flow Estimation in Rainy Scenes	Ruoteng Li*, National University of Singapore; Robby Tan, Yale-NUS College, Singapore; Loong Fah Cheong, NUS
P-3C-53	Image Generation from Sketch Constraint Using Contextual GAN	Yongyi Lu*, HKUST; Shangzhe Wu, HKUST; Yu-Wing Tai, Tencent YouTu; Chi-Keung Tang, Hong Kong University of Science and Technology
P-3C-54	Accurate Scene Text Detection through Border Semantics Awareness and Bootstrapping	Chuhui Xue, Nanyang Technological University; Shijian Lu*, Nanyang Technological University; Fangneng Zhan, Nanyang Technological University
P-3C-55	CNN-PS: CNN-based Photometric Stereo for General Non-Convex Surfaces	Satoshi Ikehata*, National Institute of Informatics
P-3C-56	Making Deep Heatmaps Robust to Partial Occlusions for 3D Object Pose Estimation	Markus Oberweger*, TU Graz; Mahdi Rad, TU Graz; Vincent Lepetit, TU Graz
P-3C-57	Recognition in Terra Incognita	Sara Beery*, Caltech
P-3C-58	Super-Resolution and Sparse View CT Reconstruction	Guangming Zang, KAUST; Ramzi Idoughi, KAUST; Mohamed Aly, KAUST; Peter Wonka, KAUST; Wolfgang Heidrich*, KAUST

P-3C-59	Modeling Visual Context is Key to Augmenting Object Detection Datasets	NIKITA DVORNIK*, INRIA; Julien Mairal, INRIA; Cordelia Schmid, INRIA
P-3C-60	Occlusions, Motion and Depth Boundaries with a Generic Network for Optical Flow, Disparity, or Scene Flow Estimation	Eddy Ilg*, University of Freiburg; Tonmoy Saikia, University of Freiburg; Margret Keuper, University of Mannheim; Thomas Brox, University of Freiburg
P-3C-61	Unsupervised Domain Adaptation for 3D Keypoint Estimation via View Consistency	Xingyi Zhou, The University of Texas at Austin; Arjun Karpur, The University of Texas at Austin; Chuang Gan, MIT; Linjie Luo, Snap Inc; Qixing Huang*, The University of Texas at Austin
P-3C-62	Improving DNN Robustness to Adversarial Attacks using Jacobian Regularization	Daniel Jakubovitz*, Tel Aviv University; Raja Giryes, Tel Aviv University
P-3C-63	A Framework for Evaluating 6-DOF Object Trackers	Mathieu Garon, Université Laval; Denis Laurendeau, Laval University; Jean-Francois Lalonde*, Université Laval
P-3C-64	Self-Supervised Relative Depth Learning for Urban Scene Understanding	Huaizu Jiang*, UMass Amherst; Erik Learned-Miller, University of Massachusetts, Amherst; Gustav Larsson, University of Chicago; Michael Maire, Toyota Technological Institute at Chicago; Greg Shakhnarovich, Toyota Technological Institute at Chicago
P-3C-65	Actor-centric Relation Network	Chen Sun*, Google; Abhinav Shrivastava, UMD / Google; Carl Vondrick, MIT; Kevin Murphy, Google; Rahul Sukthankar, Google; Cordelia Schmid, Google
P-3C-66	Self-produced Guidance for Weakly-supervised Object Localization	Xiaolin Zhang*, University of Technology Sydney; Yunchao Wei, UIUC; Guoliang Kang, UTS; Yi Yang, UTS; Thomas Huang, UIUC

P-3C-67	Attribute-Guided Face Generation Using Conditional CycleGAN	Yongyi Lu*, HKUST; Yu-Wing Tai, Tencent YouTu; Chi-Keung Tang, Hong Kong University of Science and Technology
P-3C-68	Neural Network Encapsulation	Hongyang Li*, Chinese University of Hong Kong; Bo Dai, the Chinese University of Hong Kong; Wanli Ouyang, CUHK; Xiaoyang Guo, The Chinese University of Hong Kong; Xiaogang Wang, Chinese University of Hong Kong, Hong Kong
P-3C-69	Deep Regionlets for Object Detection	Hongyu Xu*, University of Maryland; Xutao Lv, Intellifusion; Xiaoyu Wang, -; Zhou Ren, Snap Inc.; Navaneeth Bodla, University of Maryland; Rama Chellappa, University of Maryland
P-3C-70	Deep Adversarial Attention Alignment for Unsupervised Domain Adaptation: the Benefit of Target Expectation Maximization	Guoliang Kang*, UTS; Liang Zheng, Singapore University of Technology and Design; Yan Yan, UTS; Yi Yang, UTS
P-3C-71	Fighting Fake News: Image Splice Detection via Learned Self-Consistency	Jacob Huh*, Carnegie Mellon University; Andrew Liu, University of California, Berkeley; Andrew Owens, UC Berkeley; Alexei Efros, UC Berkeley
P-3C-72	Learning Monocular Depth by Distilling Cross-domain Stereo Networks	Xiaoyang Guo*, The Chinese University of Hong Kong; Hongsheng Li, Chinese University of Hong Kong; Shuai Yi, The Chinese University of Hong Kong; Jimmy Ren, Sensetime Research; Xiaogang Wang, Chinese University of Hong Kong, Hong Kong
P-3C-73	Riemannian Walk for Incremental Learning: Understanding Forgetting and Intransigence	Arslan Chaudhry*, University of Oxford; Puneet Dokania, University of Oxford; Thalaiyasingam Ajanthan, University of Oxford; Philip Torr, University of Oxford

P-3C-74	Weakly Supervised Region Proposal Network and Object Detection	Peng Tang*, Huazhong University of Science and Technology; Xinggang Wang, Huazhong Univ. of Science and Technology; Angtian Wang, Huazhong University of Science and Technology ; Yongluan Yan, Huazhong University of Science and Technology ; Wenyu Liu, Huazhong University of Science and Technology; Junzhou Huang, Tencent AI Lab; Alan Yuille, Johns Hopkins University

Poster-number	Title	Authors
P-4A-01	Viewpoint Estimation - Insights & Model	Gilad Divon, Technion; Ayellet Tal*, Technion
P-4A-02	Towards Realistic Predictors	Pei Wang*, UC San Diego; Nuno Vasconcelos, UC San Diego
P-4A-03	Group Normalization	Yuxin Wu, Facebook; Kaiming He*, Facebook Inc., USA
P-4A-04	Deep Expander Networks: Efficient Deep Networks from Graph Theory	Ameya Prabhu*, IIIT Hyderabad; Girish Varma, IIIT Hyderabad; Anoop Namboodiri, IIIT Hyderbad
P-4A-05	Learning SO(3) Equivariant Representations with Spherical CNNs	Carlos Esteves*, University of Pennsylvania; Kostas Daniilidis, University of Pennsylvania; Ameesh Makadia, Google Research; Christine Allec-Blanchette, University of Pennsylvania
P-4A-06	Video Re-localization via Cross Gated Bilinear Matching	Yang Feng*, University of Rochester; Lin Ma, Tencent AI Lab; Wei Liu, Tencent AI Lab; Tong Zhang, Tecent AI Lab; Jiebo Luo, U. Rochester
P-4A-07	A Deeply-initialized Coarse-to-fine Ensemble of Regression Trees for Face Alignment	Roberto Valle*, Universidad Politécnica de Madrid; José Buenaposada, Universidad Rey Juan Carlos; Antonio Valdés, Universidad Complutense de Madrid; Luis Baumela, Universidad Politecnica de Madrid
P-4A-08	Deep Kalman Filtering Network for Video Compression Artifact Reduction	Guo Lu*, Shanghai Jiao Tong University; Wanli Ouyang, CUHK; Dong Xu, University of Sydney; Xiaoyun Zhang, Shanghai Jiao Tong University; Zhiyong Gao, Shanghai Jiao Tong University; Ming Ting Sun, -
P-4A-09	Exploring Visual Relationship for Image Captioning	Ting Yao*, Microsoft Research; Yingwei Pan, University of Science and Technology of China; Yehao Li, Sun Yat-Sen University; Tao Mei, JD.com
P-4A-10	Sequential Clique Optimization for Video Object Segmentation	Yeong Jun Koh*, Korea University; Young-Yoon Lee, Samsung; Chang-Su Kim, Korea university

P-4A-11	Spatial Pyramid Calibration for Image Classification	Yan Wang, Shanghai Jiao Tong University; Lingxi Xie*, JHU; Siyuan Qiao, Johns Hopkins University; Ya Zhang, Cooperative Medianet Innovation Center, Shang hai Jiao Tong University; Wenjun Zhang, Shanghai Jiao Tong University; Alan Yuille, Johns Hopkins University
P-4A-12	Visual Text Correction	Amir Mazaheri*, University of Central Florida; Mubarak Shah, University of Central Florida
P-4A-13	X-ray Computational Tomography Through Scatter	Adam Geva*, Technion; Schechner Yoav, Technion; Jonathan Chernyak, Technion; Rajiv Gupta, MGH Harvard
P-4A-14	Graph Distillation for Action Detection with Privileged Information in RGB-D Videos	Zelun Luo*, Stanford University; Lu Jiang, Google; Jun-Ting Hsieh, Stanford University; Juan Carlos Niebles, Stanford University; Li Fei-Fei, Stanford University
P-4A-15	Modular Generative Adversarial Networks	Bo Zhao*, University of British Columbia; Bo Chang, University of British Columbia; Zequn Jie, Tencent AI Lab; Leonid Sigal, University of British Columbia
P-4A-16	R2P2: A ReparameteRized Pushforward Policy for Diverse, Precise Generative Path Forecasting	Nicholas Rhinehart*, CMU; Kris Kitani, CMU; Paul Vernaza, NEC Labs America
P-4A-17	DFT-based Transformation Invariant Pooling Layer for Visual Classification	Jongbin Ryu*, Hanyang University; Ming-Hsuan Yang, University of California at Merced; Jongwoo Lim, Hanyang University
P-4A-18	X2Face: A network for controlling face generation by using images, audio, and pose codes	Olivia Wiles*, University of Oxford; A Koepke, University of Oxford; Andrew Zisserman, University of Oxford
P-4A-19	Compositional Learning of Human Object Interactions	Keizo Kato, CMU; Yin Li*, CMU; Abhinav Gupta, CMU

P-4A-20	Learning to Navigate for Fine-grained Classification	Ze Yang*, Peking University; Tiange Luo, Peking University; Dong Wang, Peking University; Zhiqiang Hu, Peking University; Jun Gao, Peking University; Liwei Wang, Peking University
P-4A-21	Cross-Modal Ranking with Soft Consistency and Noisy Labels for Robust RGB-T Tracking	Chenglong Li, Anhui University; Chengli Zhu, Anhui University; Yan Huang, Institute of Automation, Chinese Academy of Sciences; Jin Tang, Anhui University; Liang Wang*, NLPR, China
P-4A-22	Light-weight CNN Architecture Design for Fast Inference	Ningning Ma*, Tsinghua; Xiangyu Zhang, Megvii Inc; Hai-Tao Zheng, Tsinghua University; Jian Sun, Megvii, Face++
P-4A-23	Fully Motion-Aware Network for Video Object Detection	Shiyao Wang*, Tsinghua University; Yucong Zhou, Beihang University; Junjie Yan, Sensetime Group Limited
P-4A-24	Shift-Net: Image Inpainting via Deep Feature Rearrangement	Zhaoyi Yan, Harbin Institute of Technology; Xiaoming Li, Harbin Institute of Technology; Mu LI, The Hong Kong Polytechnic University; Wangmeng Zuo*, Harbin Institute of Technology, China; Shiguang Shan, Chinese Academy of Sciences
P-4A-25	Choose Your Neuron: Incorporating Domain Knowledge through Neuron Importance	Ramprasaath Ramasamy Selvaraju*, Virginia Tech; Prithvijit Chattopadhyay, Georgia Institute of Technology; Mohamed Elhoseiny, Facebook; Tilak Sharma, Facebook; Dhruv Batra, Georgia Tech & Facebook AI Research; Devi Parikh, Georgia Tech & Facebook AI Research; Stefan Lee, Georgia Institute of Technology
P-4A-26	Joint 3D tracking of a deformable object in interaction with a hand	Aggeliki Tsoli*, FORTH; Antonis Argyros, CSD-UOC and ICS-FORTH
P-4A-27	Interpolating Convolutional Neural Networks Using Batch Normalization	Gratianus Wesley Putra Data*, University of Oxford; Kirjon Ngu, University of Oxford; David Murray, University of Oxford; Victor Prisacariu, University of Oxford

P-4A-28	Learning Warped Guidance for Blind Face Restoration	Xiaoming Li, Harbin Institute of Technology; Ming Liu, Harbin Institute of Technology; Yuting Ye, Harbin Institute of Technology; Wangmeng Zuo*, Harbin Institute of Technology, China; Liang Lin, Sun Yat-sen University; Ruigang Yang, University of Kentucky, USA
P-4A-29	Separable Cross-Domain Translation	Yedid Hoshen*, Facebook AI Research (FAIR); Lior Wolf, Tel Aviv University, Israel
P-4A-30	Task-driven Webpage Saliency	Quanlong Zheng*, City University of HongKong; Jianbo Jiao, City University of Hong Kong; Ying Cao, City University of Hong Kong; Rynson Lau, City University of Hong Kong
P-4A-31	Appearance-Based Gaze Estimation via Evaluation-Guided Asymmetric Regression	Yihua Cheng, Beihang University; Feng Lu*, U. Tokyo; Xucong Zhang, Max Planck Institute for Informatics and Saarland University
P-4A-32	Pivot Correlational Neural Network for Multimodal Video Categorization	Sunghun Kang*, KAIST; Junyeong Kim, KAIST; Hyunsoo Choi, SAMSUNG ELECTRONICS CO.,LTD; Sungjin Kim, SAMSUNG ELECTRONICS CO.,LTD; Chang D. Yoo, KAIST
P-4A-33	Interactive Boundary Prediction for Object Selection	Hoang Le, Portland State University; Long Mai*, Adobe Research; Brian Price, Adobe; Scott Cohen, Adobe Research; Hailin Jin, Adobe Research; Feng Liu, Portland State University
P-4A-34	Scenes-Objects-Actions: A Multi-Task, Multi-Label Video Dataset	Heng Wang*, Facebook Inc; Lorenzo Torresani, Dartmouth College; Matt Feiszli, Facebook Research; Manohar Paluri, Facebook; Du Tran, Facebook; Jamie Ray, Facebook Research; Yufei Wang, Facebook
P-4A-35	Transferable Adversarial Perturbations	Bruce Hou*, Tencent; Wen Zhou, Tencent

P-4A-36	Incremental Non-Rigid Structure-from-Motion with Unknown Focal Length	Thomas Probst, ETH Zurich; Danda Pani Paudel*, ETH Zürich; Ajad Chhatkuli , ETHZ; Luc Van Gool, ETH Zurich
P-4A-37	Semantically Aware Urban 3D Reconstruction with Plane-Based Regularization	Thomas Holzmann*, Graz University of Technology; Michael Maurer, Graz University of Technology; Friedrich Fraundorfer, Graz University of Technology; Horst Bischof, Graz University of Technology
P-4A-38	Learning to Dodge A Bullet	shi jin*, ShanghaiTech University; Jinwei Ye, Louisiana State University; Yu Ji, Plex-VR; RUIYANG LIU, ShanghaiTech University; Jingyi Yu, Shanghai Tech University
P-4A-39	Training Binary Weight Networks via Semi-Binary Decomposition	Qinghao Hu*, Institute of Automation, Chinese Academy of Sciences; Gang Li, Institute of Automation, Chinese Academy of Sciences; Peisong Wang, Institute of Automation, Chinese Academy of Sciences; yifan zhang, Institute of Automation,Chinese Academy of Sciences; Jian Cheng, Chinese Academy of Sciences, China
P-4A-40	Learnable PINs: Cross-Modal Embeddings for Person Identity	Samuel Albanie*, University of Oxford; Arsha Nagrani, Oxford University ; Andrew Zisserman, University of Oxford
P-4A-41	Toward Characteristic-Preserving Image-based Virtual Try-On Network	Bochao Wang, Sun Yet-sen University; Huabin Zheng, Sun Yat-Sen University; Xiaodan Liang*, Carnegie Mellon University; Yimin Chen, sensetime; Liang Lin, Sun Yat-sen University
P-4A-42	Deep Feature Factorization For Unsupervised Concept Discovery	Edo Collins*, EPFL; Radhakrishna Achanta, EPFL; Sabine Süsstrunk, EPFL
P-4A-43	SOD-MTGAN: Small Object Detection via Multi-Task Generative Adversarial Network	Yongqiang Zhang*, Harbin institute of Technology/KAUST; Yancheng Bai, KAUST/ISCAS; Mingli Ding, Harbin institute of Technology; Bernard Ghanem, KAUST

P-4A-44	Human Motion Analysis with Deep Metric Learning	HUSEYIN COSKUN*, Technical University of Munich; David Joseph Tan, CAMP, TU Munich; Sailesh Conjeti, Technical University of Munich; Nassir Navab, TU Munich, Germany; Federico Tombari, Technical University of Munich, Germany
P-4A-45	Dist-GAN: An Improved GAN using Distance Constraints	Ngoc-Trung Tran*, Singapore University of Technology and Design; Tuan Anh Bui, Singapore University of Technology and Design; Ngai-Man Cheung, Singapore University of Technology and Design
P-4A-46	Cross-Modal and Hierarchical Modeling of Video and Text	Bowen Zhang*, University of Southern California; Hexiang Hu, University of Southern California; Fei Sha, USC
P-4A-47	Deep Image Demosaicking using a Cascade of Convolutional Residual Denoising Networks	Filippos Kokkinos*, Skolkovo Institute of Science and Technology; Stamatis Lefkimmiatis, Skolkovo Institute of Science and Technology
P-4A-48	Deep Clustering for Unsupervised Learning of Visual Features	Mathilde Caron*, Facebook Artificial Intelligence Research; Piotr Bojanowski, Facebook; Armand Joulin, Facebook AI Research; Matthijs Douze, Facebook AI Research
P-4A-49	Domain Adaptation through Synthesis for Unsupervised Person Re-identification	Slawomir Bak*, Argo AI; Jean-Francois Lalonde, Université Laval; Pete Carr, Argo AI
P-4A-50	Facial Expression Recognition with Inconsistently Annotated Datasets	Jiabei Zeng*, Institute of Computing Technology, Chinese Academy on Sciences; Shiguang Shan, Chinese Academy of Sciences; Chen Xilin, Institute of Computing Technology, Chinese Academy of Sciences

P-4A-51	Single Shot Scene Text Retrieval	Lluis Gomez*, Universitat Autónoma de Barcelona; Andres Mafla, Computer Vision Center; Marçal Rossinyol, Universitat Autónoma de Barcelona; Dimosthenis Karatzas, Computer Vision Centre
P-4A-52	DeepVS: A Deep Learning Based Video Saliency Prediction Approach	Lai Jiang, BUAA; Mai Xu*, BUAA; Minglang Qiao, BUAA; Zulin Wang, BUAA
P-4A-53	Generalizing A Person Retrieval Model Hetero- and Homogeneously	Zhun Zhong*, Xiamen University; Liang Zheng, Singapore University of Technology and Design; Shaozi Li, Xiamen University, China; Yi Yang, University of Technology, Sydney
P-4A-54	A New Large Scale Dynamic Texture Dataset with Application to ConvNet Understanding	Isma Hadji*, York University; Rick Wildes, York University
P-4A-55	Deep Cross-modality Adaptation via Semantics Preserving Adversarial Learning for Sketch-based 3D Shape Retrieval	Jiaxin Chen, New York University Abu Dhabi; Yi Fang*, New York University
P-4A-56	BiSeNet: Bilateral Segmentation Network for Real-time Semantic Segmentation	Changqian Yu*, Huazhong University of Science and Technology; Jingbo Wang, Peking University; Chao Peng, Megvii(Face++) Inc; Changxin Gao, Huazhong University of Science and Technology; Gang Yu, Face++; Nong Sang, School of Automation, Huazhong University of Science and Technology
P-4A-57	Face De-spoofing	Yaojie Liu*, Michigan State University; Amin Jourabloo, Michigan State University; Xiaoming Liu, Michigan State University

P-4A-58	Towards End-to-End License Plate Detection and Recognition: A Large Dataset and Baseline	Zhenbo Xu*, University of Science and Technology in China; Wei Yang, University of Science and Technology in China; Ajin Meng, University of Science and Technology in China; Nanxue Lu, University of Science and Technology in China; Huan Huang, Xingtai Financial Holdings Group Co., Ltd.
P-4A-59	Self-supervised Tracking by Colorization	Carl Vondrick*, MIT; Abhinav Shrivastava, UMD / Google; Alireza Fathi, Google; Sergio Guadarrama, Google; Kevin Murphy, Google
P-4A-60	Pose Proposal Networks	Taiki Sekii*, Konica Minolta, inc.
P-4A-61	Incremental Multi-graph Matching via Diversity and Randomness based Graph Clustering	Tianshu Yu*, Arizona State University; Junchi Yan, Shanghai Jiao Tong University; baoxin Li, Arizona State University; Wei Liu, Tencent AI Lab
P-4A-62	Single Image Intrinsic Decomposition Without a Single Intrinsic Image	Wei-Chiu Ma*, MIT; Hang Chu, University of Toronto; Bolei Zhou, MIT; Raquel Urtasun, University of Toronto; Antonio Torralba, MIT
P-4A-63	Triplet Loss with Theoretical Analysis in Siamese Network for Real-Time Object Tracking	Xingping Dong, Beijing Institute of Technology; Jianbing Shen*, Beijing Institute of Technology
P-4A-64	Learning to Learn Parameterized Image Operators	Qingnan Fan, Shandong University; Dongdong Chen*, university of science and technology of china; Lu Yuan, Microsoft Research Asia; Gang Hua, Microsoft Cloud and AI; Nenghai Yu, University of Science and Technology of China; Baoquan Chen, Shandong University

P-4A-65	HBE: Hand Branch Ensemble network for real time 3D hand pose estimation	Yidan Zhou, Dalian University of Technology; Jian Lu, Laboratory of Advanced Design and Intelligent Computing, Dalian University; Kuo Du, Dalian University of Technology; Xiangbo Lin*, Dalian University of Technology; Yi Sun, Dalian University of Technology; Xiaohong Ma, Dalian University of Technology
P-4A-66	Generative Semantic Manipulation with Mask-Contrasting GAN	Xiaodan Liang*, Carnegie Mellon University
P-4A-67	Learning to Fuse Proposals from Multiple Scanline Optimizations in Semi-Global Matching	Johannes Schoenberger*, ETH Zurich; Sudipta Sinha, Microsoft Research; Marc Pollefeys, ETH Zurich
P-4A-68	Less is More: Picking Informative Frames for Video Captioning	Yangyu Chen*, University of Chinese Academy of Sciences; Shuhui Wang, vipl,ict,Chinese academic of science; Weigang Zhang, Harbin Institute of Technology, Weihai; Qingming Huang, University of Chinese Academy of Sciences, China
P-4A-69	Deep Pictorial Gaze Estimation	Seonwook Park*, ETH Zurich; Adrian Spurr, ETH Zurich; Otmar Hilliges, ETH Zurich
P-4A-70	SkipNet: Learning Dynamic Execution in Residual Networks	Xin Wang*, UC Berkeley; Fisher Yu, UC Berkeley; Zi-Yi Dou, Nanjing University; Trevor Darrell, UC Berkeley; Joseph Gonzalez, UC Berkeley
P-4A-71	Mask TextSpotter: An End-to-End Trainable Neural Network for Spotting Text with Arbitrary Shapes	Pengyuan Lyu*, Huazhong University of Science and Technology; Minghui Liao, Huazhong University of Science and Technology; Cong Yao, Megvii; Wenhao Wu, Megvii; Xiang Bai, Huazhong University of Science and Technology

P-4A-72	Deep Adaptive Attention for Joint Facial Action Unit Detection and Face Alignment	Zhiwen Shao*, Shanghai Jiao Tong University; Zhilei Liu, Tianjin University; Jianfei Cai, Nanyang Technological University; Lizhuang Ma, Shanghai Jiao Tong University
P-4A-73	Semantic Scene Understanding under Dense Fog with Synthetic and Real Data	Christos Sakaridis*, ETH Zurich; Dengxin Dai, ETH Zurich; Simon Hecker, ETH Zurich; Luc Van Gool, ETH Zurich
P-4A-74	RIDI: Robust IMU Double Integration	Hang Yan*, Washington University in St. Louis; Qi Shan, Zillow Group; Yasutaka Furukawa, Simon Fraser University
P-4A-75	Weakly-supervised Video Summarization using Variational Encoder-Decoder and Web Prior	Sijia Cai*, The Hong Kong Polytechnic University; Wangmeng Zuo, Harbin Institute of Technology; Larry Davis, University of Maryland; Lei Zhang, Hong Kong Polytechnic University, Hong Kong, China
P-4A-76	Transferring Common-Sense Knowledge for Object Detection	Krishna Kumar Singh*, University of California Davis; Santosh Divvala, Allen AI; Ali Farhadi, University of Washington; Yong Jae Lee, University of California, Davis
P-4A-77	Person Search in Videos with One Portrait Through Visual and Temporal Links	Qingqiu Huang*, CUHK; Wentao Liu, Sensetime; Dahua Lin, The Chinese University of Hong Kong
P-4A-78	Eliminating the Dreaded Blind Spot: Adapting 3D Object Detection and Monocular Depth Estimation to 360° Panoramic Imagery	Gregoire Payen de La Garanderie*, Durham University; Toby Breckon, Durham University; Amir Atapour-Abarghouei, Durham University
P-4A-79	Folded Recurrent Neural Networks for Future Video Prediction	Marc Oliu*, Universitat Oberta de Catalunya; Javier Selva, Universitat de Barcelona; Sergio Escalera, Computer Vision Center (UAB) & University of Barcelona,

P-4A-80	Deep Regression Tracking with Shrinkage Loss	Xiankai Lu, Shanghai Jiao Tong University; Chao Ma*, University of Adelaide; Bingbing Ni, Shanghai Jiao Tong University; Xiaokang Yang, Shanghai Jiao Tong University of China; Ian Reid, University of Adelaide, Australia; Ming-Hsuan Yang, University of California at Merced
P-4A-81	Stroke Controllable Fast Style Transfer with Adaptive Receptive Fields	Yongcheng Jing, Zhejiang University; Yang Liu, Zhejiang University; Yezhou Yang, Arizona State University; Zunlei Feng, Zhejiang University; Yizhou Yu, The University of Hong Kong; Dacheng Tao, University of Sydney; Mingli Song*, Zhejiang University
P-4A-82	Part-Aligned Bilinear Representations for Person Re-Identification	Yumin Suh, Seoul National University; Jingdong Wang, Microsoft Research; Kyoung Mu Lee*, Seoul National University
P-4A-83	Joint 3D Face Reconstruction and Dense Alignment with Position Map Regression Network	Yao Feng*, Shanghai Jiao Tong University; Fan Wu, CloudWalk Technology; Xiao-Hu Shao, Chongqing Institute of Green and Intelligent Technology,Chinese Academy of Sciences; Yan-Feng Wang, Shanghai Jiao Tong University; Xi Zhou, CloudWalk Technology
P-4A-84	Learning Efficient Single-stage Pedestrian Detection by Asymptotic Localization Fitting	Wei Liu*, National University of Defense Technology; Shengcai Liao, NLPR, Chinese Academy of Sciences, China; Weidong Hu, National University of Defence Technology; Xuezhi Liang, Center for Biometrics and Security Research & National Laboratory of Pattern Recognition Institute of Automation, Chinese Academy of Sciences; Xiao Chen, National University of Defense Technology

P-4A-85	Unsupervised Hard-Negative Mining from Videos for Object Detection	SouYoung Jin*, UMASS Amherst; Huaizu Jiang, UMass Amherst; Aruni RoyChowdhury, University of Massachusetts, Amherst; Ashish Singh, UMASS Amherst; Aditya Prasad, UMASS Amherst; Deep Chakraborty, UMASS Amherst; Erik Learned-Miller, University of Massachusetts, Amherst
P-4A-86	Focus, Segment and Erase: An Efficient Network for Multi-Label Brain Tumor Segmentation	Xuan Chen*, NUS; Jun Hao Liew, NUS; Wei Xiong, A*STAR Institute for Infocomm Research, Singapore; Chee-Kong Chui, NUS; Sim-Heng Ong, NUS
P-4A-87	Maximum Margin Metric Learning Over Discriminative Nullspace for Person Re-identification	T M Feroz Ali*, Indian Institute of Technology Bombay, Mumbai; Subhasis Chaudhuri, Indian Institute of Technology Bombay
P-4A-88	Efficient Relative Attribute Learning using Graph Neural Networks	Zihang Meng*, University of Wisconsin Madison; Nagesh Adluru , WISC; Vikas Singh, University of Wisconsin-Madison USA
P-4A-89	Object Level Visual Reasoning in Videos	Fabien Baradel, LIRIS; Natalia Neverova*, Facebook AI Research; Christian Wolf, INSA Lyon, France; Julien Mille, INSA Centre Val de Loire; Greg Mori, Simon Fraser University

P-4B-01	Deep Model-Based 6D Pose Refinement in RGB	Fabian Manhardt*, TU Munich; Wadim Kehl, Toyota Research Institute; Nassir Navab, Technische Universität München, Germany; Federico Tombari, Technical University of Munich, Germany
P-4B-02	ContextVP: Fully Context-Aware Video Prediction	Wonmin Byeon*, NVIDIA; Qin Wang, ETH Zurich; Rupesh Kumar Srivastava, NNAISENSE; Petros Koumoutsakos, ETH Zurich
P-4B-03	CornerNet: Detecting Objects as Paired Keypoints	Hei Law*, University of Michigan; Jia Deng, University of Michigan
P-4B-04	RelocNet: Continous Metric Learning Relocalisation using Neural Nets	Vassileios Balntas*, University of Oxford; Victor Prisacariu, University of Oxford; Shuda Li, University of Oxford
P-4B-05	Museum Exhibit Identification Challenge for the Supervised Domain Adaptation.	Piotr Koniusz*, Data61/CSIRO, ANU; Yusuf Tas, Data61; Hongguang Zhang, Australian National University; Mehrtash Harandi, Monash University; Fatih Porikli, ANU; Rui Zhang, University of Canberra
P-4B-06	Acquisition of Localization Confidence for Accurate Object Detection	Borui Jiang*, Peking University; Ruixuan Luo, Peking University; Jiayuan Mao, Tsinghua University; Tete Xiao, Peking University; Yuning Jiang, Megvii(Face++) Inc
P-4B-07	The Contextual Loss for Image Transformation with Non-Aligned Data	Roey Mechrez*, Technion; Itamar Talmi, Technion; Lihi Zelnik-Manor, Technion
P-4B-08	Saliency Benchmarking Made Easy: Separating Models, Maps and Metrics	Matthias Kümmerer*, University of Tübingen; Thomas Wallis, University of Tübingen; Matthias Bethge, University of Tübingen
P-4B-09	Multi-Attention Multi-Class Constraint for Fine-grained Image Recognition	Ming Sun, baidu; Yuchen Yuan, Baidu Inc.; Feng Zhou*, Baidu Research; Errui Ding, Baidu Inc.

P-4B-10	Look Before You Leap: Bridging Model-Free and Model-Based Reinforcement Learning for Planned-Ahead Vision-and-Language Navigation	Xin Wang*, University of California, Santa Barbara; Wenhan Xiong, University of California, Santa Barbara; Hongmin Wang, University of California, Santa Barbara; William Wang, UC Santa Barbara
P-4B-11	HandMap: Robust Hand Pose Estimation via Intermediate Dense Guidance Map Supervision	Xiaokun Wu*, University of Bath; Daniel Finnegan, University of Bath; Eamonn O'Neill, University of Bath; Yongliang Yang, University of Bath
P-4B-12	LSQ++: lower runtime and higher recall in multi-codebook quantization	Julieta Martinez*, University of British Columbia; Shobhit Zakhmi, University of British Columbia; Holger Hoos, University of British Columbia; Jim Little, University of British Columbia, Canada
P-4B-13	Multimodal Dual Attention Memory for Video Story Question Answering	Kyungmin Kim*, Seoul National University; Seong-Ho Choi, Seoul National University; Jin-Hwa Kim, Seoul National University; Byoung-Tak Zhang, Seoul National University
P-4B-14	Hierarchical Bilinear Pooling for Fine-Grained Visual Recognition	Chaojian Yu*, Huazhong University of Science and Technology; Qi Zheng, Huazhong University of Science and Technology; Xinyi Zhao, Huazhong University of Science and Technology; Peng Zhang, Huazhong University of Science and Technology; Xinge YOU, School of Electronic Information and Communications,Huazhong University of Science and Technology
P-4B-15	Dense Semantic and Topological Correspondence of 3D Faces without Landmarks	Zhenfeng Fan*, Chinese Academy of Sciences; hu xiyuan, The Chinese academy of science; chen chen, The Chinese academy of science; peng silong, The Chinese academy of science

P-4B-16	Real-Time Blind Video Temporal Consistency	Wei-Sheng Lai*, University of California, Merced; Jia-Bin Huang, Virginia Tech; Oliver Wang, Adobe Systems Inc; Eli Shechtman, Adobe Research, US; Ersin Yumer, Argo AI; Ming-Hsuan Yang, University of California at Merced
P-4B-17	Depth Estimation via Affinity Learned with Convolutional Spatial Propagation Network	Xinjing Cheng, Baidu; Peng Wang*, Baidu USA LLC; Ruigang Yang, University of Kentucky, USA
P-4B-18	Hierarchical Metric Learning and Matching for 2D and 3D Geometric Correspondences	Mohammed Fathy, University of Maryland College Park; Quoc-Huy Tran*, NEC Labs; Zeeshan Zia, Microsoft; Paul Vernaza, NEC Labs America; Manmohan Chandraker, NEC Labs America
P-4B-19	GridFace: Face Rectification via Learning Local Homography Transformations	Erjin Zhou*, Megvii Research
P-4B-20	Rethinking Spatiotemporal Feature Learning: Speed-Accuracy Trade-offs in Video Classification	Saining Xie*, UCSD; Chen Sun, Google; Jonathan Huang, Google; Zhuowen Tu, UC San Diego; Kevin Murphy, Google
P-4B-21	Deep Variational Metric Learning	Xudong Lin, Tsinghua University; Yueqi Duan, Tsinghua University; Qiyuan Dong, Tsinghua University; Jiwen Lu*, Tsinghua University; Jie Zhou, Tsinghua University, China
P-4B-22	Multi-Class Model Fitting by Energy Minimization and Mode-Seeking	Dániel Baráth*, MTA SZTAKI, CMP Prague; Jiri Matas, CMP CTU FEE
P-4B-23	A Unified Framework for Single-View 3D Reconstruction with Limited Pose Supervision	Guandao Yang*, Cornell University; Yin Cui, Cornell University; Bharath Hariharan, Cornell University

P-4B-24	Diverse Conditional Image Generation by Stochastic Regression with Latent Drop-Out Codes	Yang He*, MPI Informatics; Bernt Schiele, MPI; Mario Fritz, Max-Planck-Institut für Informatik
P-4B-25	Orthogonal Deep Features Decomposition for Age-Invariant Face Recognition	yitong wang, Tencent AI Lab; dihong gong, Tencent AI Lab; zheng zhou, Tencent AI Lab; xing ji, Tencent AI Lab; Hao Wang, Tencent AI Lab; Zhifeng Li*, Tencent AI Lab; Wei Liu, Tencent AI Lab; Tong Zhang, Tecent AI Lab
P-4B-26	HiDDeN: Hiding Data with Deep Networks	Jiren Zhu*, Stanford University; Russell Kaplan, Stanford University; Justin Johnson, Stanford University; Li Fei-Fei, Stanford University
P-4B-27	Learning and Matching Multi-View Descriptors for Registration of Point Clouds	Lei Zhou*, HKUST; Siyu Zhu, HKUST; Zixin Luo, HKUST; Tianwei Shen, HKUST; Runze Zhang, HKUST; Tian Fang, HKUST; Long Quan, Hong Kong University of Science and Technology
P-4B-28	Deep Burst Denoising	Clement Godard*, University College London; Kevin Matzen, Facebook; Matt Uyttendaele, Facebook
P-4B-29	On Offline Evaluation of Vision-based Driving Models	Felipe Codevilla, UAB; Antonio Lopez, CVC & UAB; Vladlen Koltun, Intel Labs; Alexey Dosovitskiy*, Intel Labs
P-4B-30	Distortion-Aware Convolutional Filters for Dense Prediction in Panoramic Images	Keisuke Tateno*, Technical University Munich; Nassir Navab, TU Munich, Germany; Federico Tombari, Technical University of Munich, Germany
P-4B-31	Salient Objects in Clutter: Bringing Salient Object Detection to the Foreground	Deng-Ping Fan, Nankai University; Jiang-Jiang Liu, Nankai University; Shanghua Gao, Nankai University; Qibin Hou, Nankai University; Ming-Ming Cheng*, Nankai University; Ali Borji, University of Central Florida
P-4B-32	Randomized Ensemble Embeddings	Hong Xuan*, The George Washington University; Robert Pless, George Washington University
P-4B-33	Conditional Prior Networks for Optical Flow	Yanchao Yang*, UCLA; Stefano Soatto, UCLA

P-4B-34	Adaptively Transforming Graph Matching	Fudong Wang, Wuhan University; Nan Xue, Wuhan University; yi-peng Zhang, Syracuse University; Xiang Bai, Huazhong University of Science and Technology; Gui-Song Xia*, Wuhan University
P-4B-35	Learning 3D shapes as multi-layered height maps using 2D convolutional neural networks	Kripasindhu Sarkar*, University of Kaiserslautern; Basavaraj Hampiholi, University of Kaiserslautern; Kiran Varanasi, German Research Center for Artificial Intelligence; Didier Stricker, DFKI
P-4B-36	ISNN - Impact Sound Neural Network for Material and Geometry Classification	Auston Sterling*, UNC Chapel Hill; Justin Wilson, UNC Chapel Hill; Sam Lowe, UNC Chapel Hill; Ming Lin, UNC Chapel Hill
P-4B-37	Visual Psychophysics for Making Face Recognition Algorithms More Explainable	Brandon RichardWebster*, University of Notre Dame; So Yon Kwon, Perceptive Automata; Samuel Anthony, Perceptive Automata; Christopher Clarizio, University of Notre Dame; Walter Scheirer, University of Notre Dame
P-4B-38	Show, Tell and Discriminate: Image Captioning by Self-retrieval with Partially Labeled Data	Xihui Liu*, The Chinese University of Hong Kong; Hongsheng Li, Chinese University of Hong Kong; Jing Shao, The Chinese University of Hong Kong; Dapeng Chen, The Chinese University of HongKong; Xiaogang Wang, Chinese University of Hong Kong, Hong Kong
P-4B-39	Using LIP to Gloss Over Faces in Single-Stage Face Detection Networks	Siqi Yang*, UQ ITEE; Arnold Wiliem, University of Queensland; Shaokang Chen, University of Queensland; Brian Lovell, University of Queensland
P-4B-40	Variational Wasserstein Clustering	Liang Mi*, Arizona State University; wen zhang, ASU; Xianfeng GU, Stony Brook University; Yalin Wang, Arizona State University
P-4B-41	ADVISE: Symbolism and External Knowledge for Decoding Advertisements	Keren Ye*, University of Pittsburgh; Adriana Kovashka, University of Pittsburgh

P-4B-42	Weakly- and Semi-Supervised, Non-Overlapping Instance Segmentation of Things and Stuff	Anurag Arnab*, University of Oxford; Philip Torr, University of Oxford; Qizhu Li, University of Oxford
P-4B-43	Broadcasting Convolutional Network for Visual Relational Reasoning	Simyung Chang, Seoul National University; John Yang, Seoul National University; Seonguk Park, Seoul National University; Nojun Kwak*, Seoul National University
P-4B-44	A Unified Framework for Multi-View Multi-Class Object Pose Estimation	Chi Li*, Johns Hopkins University; Jin Bai, Johns Hopkins University; Gregory D. Hager, The Johns Hopkins University
P-4B-45	Fast and Accurate Point Cloud Registration using Trees of Gaussian Mixtures	Benjamin Eckart*, NVIDIA; Kihwan Kim, NVIDIA; Kautz Jan, NVIDIA
P-4B-46	Teaching Machines to Understand Baseball Games: Large Scale Baseball Video Database for Multiple Video Understanding Tasks	Minho Shim, Yonsei University; KYUNGMIN KIM, Yonsei University; Young Hwi Kim, Yonsei University; Seon Joo Kim*, Yonsei Univ.
P-4B-47	Using Object Information for Spotting Text	Shitala Prasad*, NTU Singapore; Wai-Kin Adams Kong, Nanyang Technological University
P-4B-48	Deep Domain Generalization via Conditional Invariant Adversarial Networks	Ya Li, USTC; Xinmei Tian, USTC; Mingming Gong, CMU & U Pitt; Yajing Liu*, USTC; Tongliang Liu, The University of Sydney; Kun Zhang, Carnegie Mellon University; Dacheng Tao, University of Sydney
P-4B-49	On the Solvability of Viewing Graphs	Matthew Trager*, INRIA; Brian Osserman, UC Davis; Jean Ponce, Inria

P-4B-50	Learning Type-Aware Embeddings for Fashion Compatibility	Mariya Vasileva*, University of Illinois at Urbana-Champaign; Bryan Plummer, Boston University; Krishna Dusad, University of Illinois at Urbana-Champaign; Shreya Rajpal, University of Illinois at Urbana-Champaign; David Forsyth, Univeristy of Illinois at Urbana-Champaign; Ranjitha Kumar, UIUC: CS
P-4B-51	Visual Coreference Resolution in Visual Dialog using Neural Module Networks	Satwik Kottur*, Carnegie Mellon University; José M. F. Moura, Carnegie Mellon University; Devi Parikh, Georgia Tech & Facebook AI Research; Dhruv Batra, Georgia Tech & Facebook AI Research; Marcus Rohrbach, Facebook AI Research
P-4B-52	Hard-Aware Point-to-Set Deep Metric for Person Re-identification	Rui Yu*, Huazhong University of Science and Technology; Zhiyong Dou, Huazhong University of Science and Technology; Song Bai, HUST; ZHAOXIANG ZHANG, Chinese Academy of Sciences, China; Yongchao Xu, HUST; Xiang Bai, Huazhong University of Science and Technology
P-4B-53	Gray box adversarial training	Vivek B S*, Indian Institute of Science; Konda Reddy Mopuri, Indian Institute of Science, Bangalore; Venkatesh Babu RADHAKRISHNAN, Indian Institute of Science
P-4B-54	Exploiting Vector Fields for Geometric Rectification of Distorted Document Images	Gaofeng Meng*, Chinese Academy of Sciences; Yuanqi Su, Xi'an Jiaotong University; Ying Wu, Northwestern University; SHIMING XIANG, Chinese Academy of Sciences, China; Chunhong Pan, Institute of Automation, Chinese Academy of Sciences
P-4B-55	Revisiting RCNN: On Awakening the Classification Power of Faster RCNN	Yunchao Wei*, UIUC; Bowen Cheng, UIUC; Honghui Shi, UIUC; Rogerio Feris, IBM Research; Jinjun Xiong, IBM Thomas J. Watson Research Center; Thomas Huang, UIUC

P-4B-56	DeepTAM: Deep Tracking and Mapping	Huizhong Zhou*, University of Freiburg; Benjamin Ummenhofer, University of Freiburg; Thomas Brox, University of Freiburg
P-4B-57	On Regularized Losses for Weakly-supervised CNN Segmentation	Meng Tang*, University of Waterloo; Ismail Ben Ayed, ETS; Federico Perazzi, Disney Research; Abdelaziz Djelouah, Disney Research; Christopher Schroers, Disney Research; Yuri Boykov, University of Waterloo
P-4B-58	ShapeCodes: Self-Supervised Feature Learning by Lifting Views to Viewgrids	Dinesh Jayaraman*, UC Berkeley; Ruohan Gao, University of Texas at Austin; Kristen Grauman, University of Texas
P-4B-59	A Minimal Closed-Form Solution for Multi-Perspective Pose Estimation using Points and Lines	Pedro Miraldo*, Instituto Superior Técnico, Lisboa; Tiago Dias, Institute for systems and robotics; Srikumar Ramalingam, University of Utah
P-4B-60	Interaction-aware Spatio-temporal Pyramid Attention Networks for Action Classification	Yang Du, NLPR; Chunfeng Yuan*, NLPR; Weiming Hu, Institute of Automation,Chinese Academy of Sciences
P-4B-61	Towards Privacy-Preserving Visual Recognition via Adversarial Training: A Pilot Study	Zhenyu Wu, Texas A&M University; Zhangyang Wang*, Texas A&M University; Zhaowen Wang, Adobe Research; Hailin Jin, Adobe Research
P-4B-62	Polarimetric Three-View Geometry	Lixiong Chen, National Institute of Informatics; Yinqiang Zheng*, National Institute of Informatics; Art Subpa-asa, Tokyo Institute of Technology; Imari Sato, National Institute of Informatics
P-4B-63	SketchyScene: Richly-Annotated Scene Sketches	Changqing Zou*, University of Maryland (UMD); Qian Yu, Queen Mary University of London; Ruofei Du, UMD; Haoran Mo, sun yat sen university; Yi-Zhe Song, Queen Mary University of London; Tao Xiang, Queen Mary, University of London, UK; Chengying Gao, sun yat sen university; Baoquan Chen, Shandong University; Hao Zhang, SFU

P-4B-64	Bi-Real Net: Enhancing the Performance of 1-bit CNNs with Improved Representational Capability and Advanced Training Algorithm	zechun liu*, HKUST; Baoyuan Wu, Tencent AI Lab; Wenhan Luo, Tencent AI Lab; Xin Yang, Huazhong University of Science and Technology; Wei Liu, Tencent AI Lab; Kwang-Ting Cheng, Hong Kong University of Science and Technology
P-4B-65	Deep Continuous Fusion for Multi-Sensor 3D Object Detection	Ming Liang*, Uber; Shenlong Wang, Uber ATG, University of Toronto; Bin Yang, Uber ATG, University of Toronto; Raquel Urtasun, Uber ATG
P-4B-66	Focus on the Hard Things: Dynamic Task Prioritization for Multitask Learning	Michelle Guo*, Stanford University; Albert Haque, Stanford University; De-An Huang, Stanford University; Serena Yeung, Stanford University; Li Fei-Fei, Stanford University
P-4B-67	Domain transfer through deep activation matching	Haoshuo Huang*, Tsinghua University; Qixing Huang, The University of Texas at Austin; Philipp Kraehenbuehl, UT Austin
P-4B-68	Joint Blind Motion Deblurring and Depth Estimation of Light Field	Dongwoo Lee, Seoul Ntional University; Haesol Park, Seoul National University; In Kyu Park, Inha University; Kyoung Mu Lee*, Seoul National University
P-4B-69	Learning to Look around Objects for Top-View Representations of Outdoor Scenes	Samuel Schulter*, NEC Labs; Menghua Zhai, University of Kentucky; Nathan Jacobs, University of Kentucky; Manmohan Chandraker, NEC Labs America
P-4B-70	Data-Driven Sparse Structure Selection for Deep Neural Networks	Zehao Huang*, TuSimple; Naiyan Wang, TuSimple
P-4B-71	Reconstruction-based Pairwise Depth Dataset for Depth Image Enhancement Using CNN	Junho Jeon, POSTECH; Seungyong Lee*, POSTECH
P-4B-72	A Geometric Perspective on Structured Light Coding	Mohit Gupta*, University of Wisconsin-Madison, USA ; Nikhil Nakhate, University of Wisconsin-Madison

P-4B-73	3D Ego-Pose Estimation via Imitation Learning	Ye Yuan*, Carnegie Mellon University; Kris Kitani, CMU
P-4B-74	Unsupervised Learning of Multi-Frame Optical Flow with Occlusions	Joel Janai*, Max Planck Institute for Intelligent Systems; Fatma Güney, University of Oxford; Anurag Ranjan, MPI for Intelligent Systems; Michael Black, Max Planck Institute for Intelligent Systems; Andreas Geiger, MPI-IS and University of Tuebingen
P-4B-75	Dynamic Conditional Networks for Few-Shot Learning	Fang Zhao, National University of Singapore; Jian Zhao*, National University of Singapore; Yan Shuicheng, National University of Singapore; Jiashi Feng, NUS
P-4B-76	3DFeat-Net: Weakly Supervised Local 3D Features for Rigid Point Cloud Registration	Zi Jian Yew*, National University of Singapore; Gim Hee Lee, National University of SIngapore
P-4B-77	Learning to Forecast and Refine Residual Motion for Image-to-Video Generation	Long Zhao*, Rutgers University; Xi Peng, Rutgers University; Yu Tian, Rutgers; Mubbasir Kapadia, Rutgers; Dimitris Metaxas, Rutgers
P-4B-78	Learn-to-Score: Efficient 3D Scene Exploration by Predicting View Utility	Benjamin Hepp*, ETH Zurich; Debadeepta Dey, Microsoft; Sudipta Sinha, Microsoft Research; Ashish Kapoor, Microsoft; Neel Joshi, -; Otmar Hilliges, ETH Zurich
P-4B-79	Deep Co-Training for Semi-Supervised Image Recognition	Siyuan Qiao*, Johns Hopkins University; Wei Shen, Shanghai University; Zhishuai Zhang, Johns Hopkins University; Bo Wang, Hikvision Research Institue; Alan Yuille, Johns Hopkins University
P-4B-80	Attention-aware Deep Adversarial Hashing for Cross Modal Retrieval	Xi Zhang, Sun Yat-Sen University; Hanjiang Lai*, Sun Yat-Sen university; Jiashi Feng, NUS

P-4B-81	Remote Photoplethysmography Correspondence Feature for 3D Mask Face Presentation Attack Detection	Siqi Liu*, Department of Computer Science, Hong Kong Baptist University; Xiangyuan Lan, Department of Computer Science, Hong Kong Baptist University; PongChi Yuen, Department of Computer Science, Hong Kong Baptist University
P-4B-82	Semi-Supervised Generative Adversarial Hashing for Image Retrieval	Guan'an Wang*, Chinese Academy of Sciences; Qinghao Hu, Chinese Academy of Sciences; Jian Cheng, Chinese Academy of Sciences, China; Zengguang Hou, Chinese Academy of Sciences
P-4B-83	Improving Spatiotemporal Self-Supervision by Deep Reinforcement Learning	Uta Büchler*, Heidelberg University; Biagio Brattoli, Heidelberg University; Bjorn Ommer, Heidelberg University
P-4B-84	AutoLoc: Weakly-supervised Temporal Action Localization in Untrimmed Videos	Zheng Shou*, Columbia University; Hang Gao, Columbia University; Lei Zhang, Microsoft Research; Kazuyuki Miyazawa, Mitsubishi Electric; Shih-Fu Chang, Columbia University
P-4B-85	Revisiting Autofocus for Smartphone Cameras	Abdullah Abuolaim*, York University; Abhijith Punnappurath, York University; Michael Brown, York University
P-4B-86	Contour Knowledge Transfer for Salient Object Detection	Xin Li, UESTC; Fan Yang*, UESTC; Hong Cheng, UESTC; Wei Liu, Digital Media Technology Key Laboratory of Sichuan Province, UESTC; Dinggang Shen, UNC
P-4B-87	Deep Volumetric Video From Very Sparse Multi-View Performance Capture	Zeng Huang*, University of Southern California; Tianye Li, University of Southern California; Weikai Chen, USC Institute for Creative Technology; Yajie Zhao, USC Institute for Creative Technology ; Jun Xing, Institute for Creative Technologies, USC; Chloe LeGendre, USC Institute for Creative Technology ; Linjie Luo, Snap Inc; Chongyang Ma, Snap Inc.; Hao Li, Pinscreen/University of Southern California/USC ICT

P-4B-88	Person Re-identification with Deep Similarity-Guided Graph Neural Network	Yantao Shen*, The Chinese University of Hong Kong; Hongsheng Li, Chinese University of Hong Kong; Shuai Yi, The Chinese University of Hong Kong; Xiaogang Wang, Chinese University of Hong Kong, Hong Kong
P-4B-89	Deep Component Analysis via Alternating Direction Neural Networks	Calvin Murdock*, Carnegie Mellon University; MingFang Chang, Carnegie Mellon University; Simon Lucey, CMU
P-4B-90	Understanding Perceptual and Conceptual Fluency at a Large Scale	Meredith Hu*, Cornell University; Ali Borji, University of Central Florida
P-4B-91	Look Deeper into Depth: Monocular Depth Estimation with Semantic Booster and Attention-Driven Loss	Jianbo Jiao*, City University of Hong Kong; Ying Cao, City University of Hong Kong; Yibing Song, Tencent AI Lab; Rynson Lau, City University of Hong Kong

Diamond

amazon

DiDi

facebook

Google AI

Microsoft

Tencent YouTu Lab

UBER

Platinum

Gold

Silver

Adobe RESEARCH

AUTONOMOUS INTELLIGENT DRIVING

字节跳动 ByteDance

Alibaba Group 阿里巴巴集团 | FASHION

cookpad

眼控科技 EMDATA

figure eight

gauss

GENIUS PROS 光 珀 智 能

HUAWEI

IIAI Inception Institute of Artificial Intelligence

iMerit human-empowered computing

logitech

MALONG 码隆

Mapillary

MathWorks

MERANTIX

Mighty AI

nuTonomy

OMRON

sama SOURCE

SECOND SPECTRUM

SIEMENS Ingenuity for life

Snap Inc.

UNIVRSES

Valeo

依图 | YITU

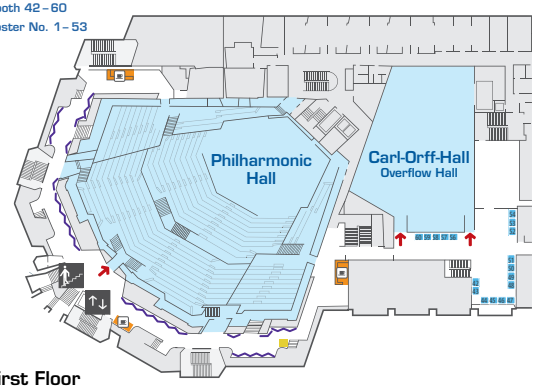
Bronze

Startup

Second Floor

Booth 42 – 60

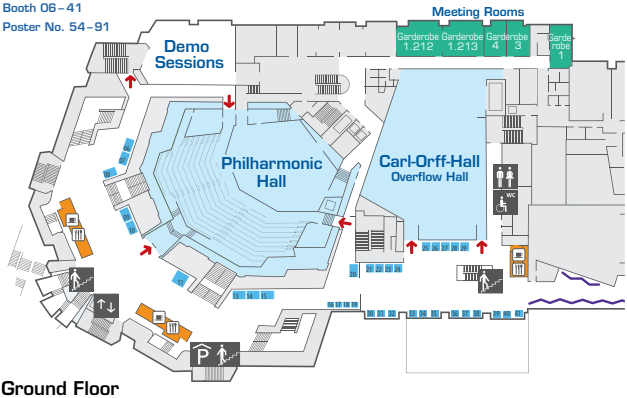
Poster No. 1 – 53



First Floor

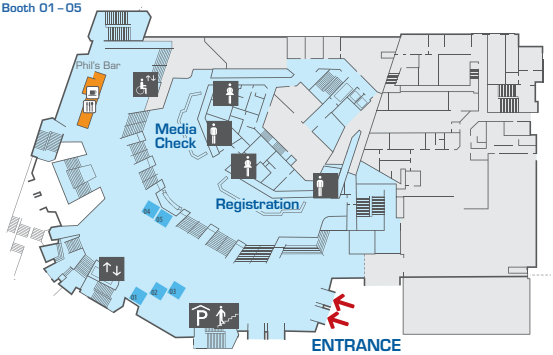
Booth 06 – 41

Poster No. 54 – 91



Ground Floor

Booth 01 – 05



Company	Booth
Alibaba FashionAI	40
Amazon	03
Apple	58
Artisense GmbH	52
Autonomous Intelligent Driving GmbH	54
BAIDU	08
BMW Group	06
Bytedance Inc.	17
Continental	09
Cookpad	37
D-ARIA	44
DREAMVU.INC.	51
Facebook	04
Fashwell	50
Gauss Surgical	43
Google	11
Hangzhou Genius Pros Technologies Co. Ltd.	56
Huawei	59
iMerit	53
Inception Institute of Artificial Intelligence	45
iRobot	14
JD.com	22
Logitech	34
LUCID Vision Labs	49
Lyft	13
Malong Technologies	39
Mapillary AB	32
MathWorks	41
Merantix	33
Microsoft Corporation	05
Mighty AI	19
MoonVision	47
MVTec Software GmbH	20
NAVER / LINE	15
NAVER LABS	10
NextAI / NEXT Canada	42
nuTonomy	46
OMRON	31
Panasonic	26
PTC	28
Robert Bosch GmbH	07
Roboception	29
Samasource	30
SAP SE	27
Second Spectrum	36
Shanghai Em-data Technology Co., Ltd.	38
SIEMENS	57
SPORTLOGiQ	21
Tobii AB	25
Uber	02
Univrses	18
Valeo	16
Watrix Technology	48
YITU Technology	60
YouTu Lab,Tencent	01
Zalando	24
ZEISS	23

Welcome Reception

Date/Time: Monday, 10 September 2018 from 06.00 pm
Location: Gasteig
Address: Rosenheimer Str. 5, 81667 Munich

Congress Dinner

Date/Time: Wednesday, 12 September 2018 from 07.30 pm
Location: Löwenbräukeller
Address: Nymphenburger Str. 2, 80335 Munich

Please note that the congress dinner is sold out. Only attendees who have signed up for the dinner in advance will gain entrance.