

## Report on ICCV 2017

The International Conference on Computer Vision (ICCV) was held on 22–29 October on beautiful Lido island of Venice, Italy. The biennial premier computer vision event was hosted at the historic Palazzo del Cinema (Venice Convention Center), recognized worldwide as the setting for the oldest film festival in the world, the Venice International Film Festival.



Outside the conference venue.

Similar to CVPR and ECCV, ICCV is considered one of the top-quality conferences in computer vision. Co-sponsored by the Institute of Electrical and Electronics Engineers (IEEE) and the Computer Vision Foundation (CVF), ICCV this year comprised the main conference, 44 co-located workshops, 60 exhibitors and 9 tutorials. It had 3107 attendees, 2143 submissions and 621 accepted papers, all highest in the long-lasting 30-year history of ICCV.

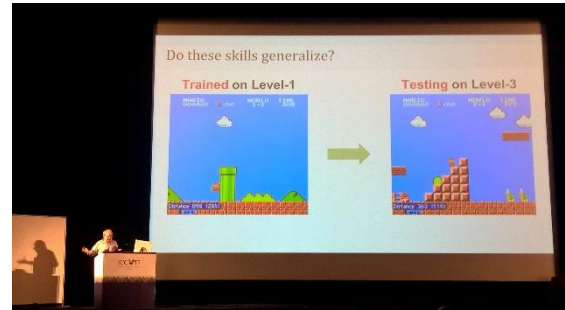
A week-long conference organised and completed with a high degree of excellence offered oral and spotlight presentations as well as poster sessions. Here, I would like to provide an overview of some of the important studies.



Michael Black on the history of human body modelling.

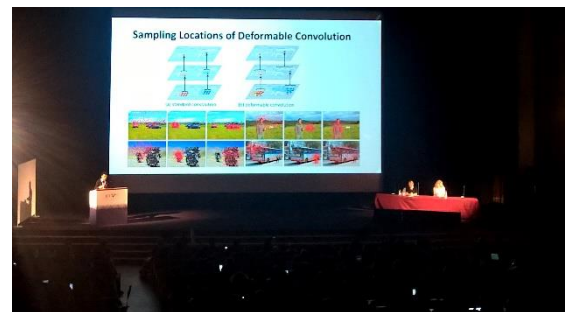
Michael J. Black from the Max Planck Institute for Intelligent Systems, where he leads the Perceiving Systems department, presented the latest developments in the field of human body modelling as an invited speaker at *PoseTrack Challenge: Human Pose Estimation and Tracking in the Wild*. A well-known figure in the research community as

well as a co-founder of Body Labs Inc., which was recently acquired by Amazon, he outlined his vision for the open problems in full-body avatars, accurate 3D human pose and shape from single images and body shape estimation from clothed scans.



Alexei Efros on 'The Revolution will not be supervised'.

Alexei A. Efros from the University of California, Berkeley gave one of his inspiring talks on reinforcement learning, aptly titled 'The Revolution will not be supervised', at the *Beyond Supervised Learning* Workshop. He also remarked on the recent study of his research group at the Berkeley Artificial Intelligence Research Lab (BAIR) that they can turn horses into zebras with the power of generative adversarial networks.



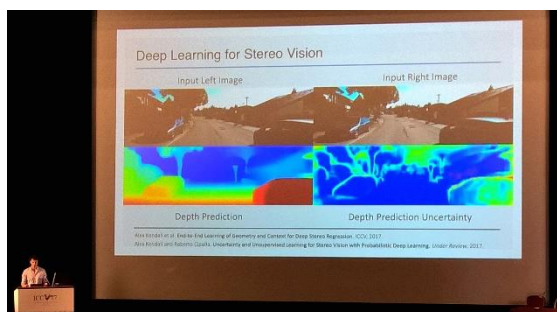
Jifeng Dai on 'Deformable Convolutional Networks'.

Jifeng Dai from Microsoft Research Asia presented an approach to model geometric transformations within Convolutional Neural Networks (CNNs). The study proposes new modules to enhance the limited capability of traditional neural networks. These new modules, called deformable convolution and deformable RoI pooling, learn 2D offsets from the target tasks without additional supervision.



Kaiming He on 'Mask R-CNN'.

This year's Best Paper Award (Marr prize) went to 'Mask R-CNN' by Kaiming He from Facebook AI Research (FAIR), along with Georgia Gkioxari, Piotr Dollár and Ross Girshick. The study offers a general framework for instance segmentation by predicting object masks in instance-level as an extension to the famous paper, 'Faster R-CNN'. Tsung-Yi Lin et al. from Facebook AI Research (FAIR), won the Best Student Paper Award with their study, 'Focal Loss for Dense Object Detection'. Tomaso Poggio, who is one of the founders of computational neuroscience, received the Azriel Rosenfeld Lifetime Achievement Award while the Distinguished Researcher Award was given to Luc van Gool and Richard Szeliski whose research has significantly contributed to the field of computer vision.



Alex Kendall on geometry meets deep learning.

Alex Kendall from the University of Cambridge focused on the understanding of classical computer vision geometry to design better deep convolutional neural network architectures at the workshop on *Geometry Meets Deep Learning*. In the invited talk, titled 'Has end-to-end deep learning killed geometry or can we do better?', he presented an overview of his work in semantic segmentation, stereo vision and multi-task learning in the domain of scene understanding.



Alex Bronstein on the correspondence problem.

Alex M. Bronstein from Technion – Israel Institute of Technology discussed several geometric notions to solve the correspondence problem efficiently at the workshop on *Multiview Relationships on 3D Data*. He showed very appealing examples on how to compute the correspondence between three-dimensional objects combined with deep learning techniques.



A packed poster session.

As an upward trend, deep learning studies were dominant throughout the conference, similar to other machine learning and computer vision conferences. In fact, *learning*, *network*, *image* and *deep* are the most used words in the title of all submitted papers at ICCV, respectively. Moreover, according to the statistics from the program committee, *video and language* and *vision for autonomous driving* have the highest acceptance rate and were seen as hot topics at ICCV 2017. It would not be wrong to say that the hype around deep learning will not be settling anytime soon.



Anil Bas explains his work to delegates.

Overall, ICCV was the perfect hub for bringing researchers from academia and industry together to share and collaborate on their ideas. More information about the conference can be found at <http://iccv2017.thecvf.com>. The next ICCV will be held in Seoul, South Korea, in 2019 and after that, it will take place in Montreal, Canada, in 2021.

Finally, I would like to thank the co-authors of our collaborative study between the University of York and the University of Surrey: Patrik Huber, William A. P. Smith, Muhammad Awais and Josef Kittler. I would also like to thank the British Machine Vision Association (BMVA) for their generous travel bursary which made it possible for me to attend this great conference.

Anil Bas  
University of York  
Email: [ab1792@york.ac.uk](mailto:ab1792@york.ac.uk)



The sun sets on Venice.