2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2018)

Salt Lake City, Utah, USA
18-22 June 2018

Pages 1-731

IEEE Catalog Number: CFP18003-POD
ISBN: 978-1-5386-6421-6

Copyright © 2018 by the Institute of Electrical and Electronics Engineers, Inc.
All Rights Reserved

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

****** This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP18003-POD
ISBN (Print-On-Demand):	978-1-5386-6421-6
ISBN (Online):	978-1-5386-6420-9
ISSN:	1063-6919

Additional Copies of This Publication Are Available From:

Curran Associates, Inc
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: (845) 758-0400
Fax: (845) 758-2633
E-mail: curran@proceedings.com
Web: www.proceedings.com

2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition

CVPR 2018

Table of Contents

Message from the General and Program Chairs ..... **cxviii**
Organizing Committee and Area Chairs ..... **cxx**
Reviewers ..... **cxxi**

Oral 1-1A

O1-1A: Object Recognition and Scene Understanding I

Embodied Question Answering 1
Abhishek Das (Georgia Institute of Technology), Samyak Datta (Georgia Institute of Technology), Georgia Gkioxari (Facebook AI Research), Stefan Lee (Georgia Institute of Technology), Devi Parikh (Facebook AI Research), Devi Parikh (Georgia Institute of Technology), Dhruv Batra (Facebook AI Research), and Dhruv Batra (Georgia Institute of Technology)

Learning by Asking Questions 11
Ishan Misra (Carnegie Mellon University), Ross Girshick (Facebook AI Research), Rob Fergus (Facebook AI Research), Martial Hebert (Carnegie Mellon University), Abhinav Gupta (Carnegie Mellon University), and Laurens van der Maaten (Carnegie Mellon University)

Oral and Spotlight 1-1B

O1-1B: Analyzing Humans in Images I

Finding Tiny Faces in the Wild with Generative Adversarial Network 21
Yancheng Bai (King Abdullah University of Science and Technology (KAUST)), Yancheng Bai (Chinese Academy of Sciences (CAS)), Yongqiang Zhang (King Abdullah University of Science and Technology (KAUST)), Yongqiang Zhang (Harbin Institute of Technology (HIT)), Mingli Ding (Harbin Institute of Technology (HIT)), and Bernard Ghanem (King Abdullah University of Science and Technology (KAUST))

Learning Face Age Progression: A Pyramid Architecture of GANs 31
Hongyu Yang (Beihang University, China), Di Huang (Beihang University, China), Yunhong Wang (Beihang University, China), and Anil K. Jain (Michigan State University, USA)

PairedCycleGAN: Asymmetric Style Transfer for Applying and Removing Makeup 40
Huiwen Chang (Princeton University), Jingwan Lu (Adobe Research), Fisher Yu (UC Berkeley), and Adam Finkelstein (Princeton University)

GANerated Hands for Real-Time 3D Hand Tracking from Monocular RGB 49
Franziska Mueller (MPI Informatics), Franziska Mueller (Saarland Informatics Campus), Florian Bernard (MPI Informatics), Florian Bernard (Saarland Informatics Campus), Oleksandr Sotnychenko (MPI Informatics), Oleksandr Sotnychenko (Saarland Informatics Campus), Dushyant Mehta (MPI Informatics), Dushyant Mehta (Saarland Informatics Campus), Srinath Sridhar (Stanford University), Dan Casas (Univ. Rey Juan Carlos), Christian Theobalt (MPI Informatics), and Christian Theobalt (Saarland Informatics Campus)

Learning Pose Specific Representations by Predicting Different Views 60
Georg Poier (Institute for Computer Graphics and Vision Graz University of Technology Austria), David Schinagl (Institute for Computer Graphics and Vision Graz University of Technology Austria), and Horst Bischof (Institute for Computer Graphics and Vision Graz University of Technology Austria)

Weakly and Semi Supervised Human Body Part Parsing via Pose-Guided Knowledge Transfer 70
Hao-Shu Fang (Shanghai Jiao Tong University, China), Guansong Lu (Shanghai Jiao Tong University, China), Xiaolin Fang (Zhejiang University, China), Jianwen Xie (University of California, Los Angeles, USA), Yu-Wing Tai (Tencent YouTu), and Cewu Lu (Shanghai Jiao Tong University, China)

S1-1B

Person Transfer GAN to Bridge Domain Gap for Person Re-identification 79
Longhui Wei (Peking University, Beijing, China), Shiliang Zhang (Peking University, Beijing, China), Wen Gao (Peking University, Beijing, China), and Qi Tian (University of Texas at San Antonio, USA)

Cross-Modal Deep Variational Hand Pose Estimation 89
Adrian Spurr (ETH Zurich), Jie Song (ETH Zurich), Seonwook Park (ETH Zurich), and Otmar Hilliges (ETH Zurich)

Disentangled Person Image Generation 99
Liqian Ma (KU-Leuven/PSI, Toyota Motor Europe (TRACE)), Qianru Sun (Max Planck Institute for Informatics, Saarland Informatics Campus), Stamatios Georgoulis (KU-Leuven/PSI, Toyota Motor Europe (TRACE)), Luc Van Gool (KU-Leuven/PSI, Toyota Motor Europe (TRACE)), Bernt Schiele (Max Planck Institute for Informatics, Saarland Informatics Campus), and Mario Fritz (Max Planck Institute for Informatics, Saarland Informatics Campus)

Super-FAN: Integrated Facial Landmark Localization and Super-Resolution of Real-World Low Resolution Faces in Arbitrary Poses with GANs 109
Adrian Bulat (The University of Nottingham, United Kingdom) and Georgios Tzimiropoulos (The University of Nottingham, United Kingdom)

Multistage Adversarial Losses for Pose-Based Human Image Synthesis 118
Chenyang Si (Center for Research on Intelligent Perception and Computing (CRIPAC), National Laboratory of Pattern Recognition (NLPR)), Chenyang Si (University of Chinese Academy of Sciences (UCAS)), Wei Wang (Center for Research on Intelligent Perception and Computing (CRIPAC), National Laboratory of Pattern Recognition (NLPR)), Wei Wang (University of Chinese Academy of Sciences (UCAS)), Liang Wang (Center for Research on Intelligent Perception and Computing (CRIPAC), National Laboratory of Pattern Recognition (NLPR)), Liang Wang (Center for Excellence in Brain Science and Intelligence Technology (CEBSIT), Institute of Automation, Chinese Academy of Sciences (CASIA)), Liang Wang (University of Chinese Academy of Sciences (UCAS)), Tieniu Tan (Center for Research on Intelligent Perception and Computing (CRIPAC), National Laboratory of Pattern Recognition (NLPR)), Tieniu Tan (Center for Excellence in Brain Science and Intelligence Technology (CEBSIT), Institute of Automation, Chinese Academy of Sciences (CASIA)), and Tieniu Tan (University of Chinese Academy of Sciences (UCAS))

Oral and Spotlight 1-1C

O1-1C: 3D Vision I

Rotation Averaging and Strong Duality 127
Anders Eriksson (Queensland University of Technology), Carl Olsson (Chalmers University of Technology), Carl Olsson (Lund University), Fredrik Kahl (Chalmers University of Technology), Fredrik Kahl (Lund University), and Tat-Jun Chin (The University of Adelaide)

Hybrid Camera Pose Estimation 136
Federico Camposeco (ETH Zürich), Andrea Cohen (ETH Zürich), Marc Pollefeys (ETH Zürich; Microsoft), and Torsten Sattler (ETH Zürich)

A Certifiably Globally Optimal Solution to the Non-minimal Relative Pose Problem 145
Jesus Briales (MAPIR-UMA Group University of Malaga, Spain), Laurent Kneip (Mobile Perception Lab SIST ShanghaiTech), and Javier Gonzalez-Jimenez (MAPIR-UMA Group University of Malaga, Spain)

S1-1C

Single View Stereo Matching 155
Yue Luo (SenseTime Research), Jimmy Ren (SenseTime Research), Mude Lin (SenseTime Research), Jiahao Pang (SenseTime Research), Wenxiu Sun (SenseTime Research), Hongsheng Li (The Chinese University of Hong Kong, Hong Kong SAR, China), Liang Lin (SenseTime Research), and Liang Lin (Sun Yat-sen University, China)

Fight Ill-Posedness with Ill-Posedness: Single-shot Variational Depth Super-Resolution from Shading 164
Bjoern Haefner (Technical University of Munich, Germany), Yvain Quéau (Technical University of Munich, Germany), Thomas Möllenhoff (Technical University of Munich, Germany), and Daniel Cremers (Technical University of Munich, Germany)

Deep Depth Completion of a Single RGB-D Image 175
Yinda Zhang (Princeton University) and Thomas Funkhouser (Princeton University)

Multi-view Harmonized Bilinear Network for 3D Object Recognition 186
Tan Yu (Nanyang Technological University), Jingjing Meng (State University of New York at Buffalo), and Junsong Yuan (State University of New York at Buffalo)

PPFNet: Global Context Aware Local Features for Robust 3D Point Matching 195
Haowen Deng (Technische Universitat München, Germany; Siemens AG, München, Germany; National University of Defense Technology, China), Tolga Birdal (Technische Universitat München, Germany), and Slobodan Ilic (Technische Universitat München, Germany)

FoldingNet: Point Cloud Auto-Encoder via Deep Grid Deformation 206
Yaoqing Yang (Carnegie Mellon University), Chen Feng (Mitsubishi Electric Research Laboratories (MERL)), Yiru Shen (Clemson University), and Dong Tian (Mitsubishi Electric Research Laboratories (MERL))

A Papier-Mache Approach to Learning 3D Surface Generation 216
Thibault Groueix (LIGM (UMR 8049), École des Ponts, UPE), Matthew Fisher (Adobe Research), Vladimir G. Kim (Adobe Research), Bryan C. Russell (Adobe Research), and Mathieu Aubry (LIGM (UMR 8049), École des Ponts, UPE)

LEGO: Learning Edge with Geometry all at Once by Watching Videos 225
Zhenheng Yang (University of Southern California), Peng Wang (Baidu Research), Yang Wang (Baidu Research), Wei Xu (National Engineering Laboratory for Deep Learning Technology and Applications), and Ram Nevatia (University of Southern California)

Poster Session P1-1

Five-Point Fundamental Matrix Estimation for Uncalibrated Cameras 235
Daniel Barath (Machine Perception Research Laboratory, MTA SZTAKI, Budapest, Hungary; Czech Technical University, Prague, Czech Republic)

PointFusion: Deep Sensor Fusion for 3D Bounding Box Estimation 244
Danfei Xu (Stanford Unviersity), Dragomir Anguelov (Zoox Inc.), and Ashesh Jain (Zoox Inc.)

Scalable Dense Non-rigid Structure-from-Motion: A Grassmannian Perspective 254
Suryansh Kumar (Australian National University), Anoop Cherian (Australian National University), Anoop Cherian (ACRV), Anoop Cherian (MERL Cambridge, MA), Yuchao Dai (Australian National University), Yuchao Dai (NPU China), Hongdong Li (Australian National University), and Hongdong Li (ACRV)

GVCNN: Group-View Convolutional Neural Networks for 3D Shape Recognition 264
Yifan Feng (Xiamen University), Zizhao Zhang (Tsinghua University), Xibin Zhao (Tsinghua University), Rongrong Ji (Xiamen University), and Yue Gao (Tsinghua University)

Depth and Transient Imaging with Compressive SPAD Array Cameras 273
Qilin Sun (KAUST), Xiong Dun (KAUST), Yifan Peng (University of British Columbia; KAUST), and Wolfgang Heidrich (KAUST)

GeoNet: Geometric Neural Network for Joint Depth and Surface Normal Estimation 283
Xiaojuan Qi (The Chinese University of Hong Kong), Renjie Liao (Uber Advanced Technologies Group; University of Toronto), Zhengzhe Liu (The Chinese University of Hong Kong), Raquel Urtasun (Uber Advanced Technologies Group; University of Toronto), and Jiaya Jia (YouTu Lab, Tencent; The Chinese University of Hong Kong)

Real-Time Seamless Single Shot 6D Object Pose Prediction 292
Bugra Tekin (EPFL), Sudipta N. Sinha (Microsoft Research), and Pascal Fua (EPFL)

Factoring Shape, Pose, and Layout from the 2D Image of a 3D Scene 302
Shubham Tulsiani (University of California, Berkeley), Saurabh Gupta (University of California, Berkeley), David Fouhey (University of California, Berkeley), Alexei A. Efrosefros (University of California, Berkeley), and Jitendra Malik (University of California, Berkeley)

Monocular Relative Depth Perception with Web Stereo Data Supervision 311
Ke Xian (Huazhong University of Science and Technology, China), Chunhua Shen (The University of Adelaide, Australia), Zhiguo Cao (Huazhong University of Science and Technology, China), Hao Lu (Huazhong University of Science and Technology, China), Yang Xiao (Huazhong University of Science and Technology, China), Ruibo Li (Huazhong University of Science and Technology, China), and Zhenbo Luo (Samsung Research Beijing, China)

Spline Error Weighting for Robust Visual-Inertial Fusion 321
Hannes Ovrén (Linköping University, Sweden) and Per-Erik Forssén (Linköping University, Sweden)

Single-Image Depth Estimation Based on Fourier Domain Analysis 330
Jae-Han Lee (Korea University), Minhyeok Heo (Korea University), Kyung-Rae Kim (Korea University), and Chang-Su Kim (Korea University)

Unsupervised Learning of Monocular Depth Estimation and Visual Odometry with Deep Feature Reconstruction 340
Huangying Zhan (The University of Adelaide), Huangying Zhan (Australian Centre for Robotic Vision), Ravi Garg (The University of Adelaide; Australian Centre for Robotic Vision), Chamara Saroj Weerasekera (The University of Adelaide; Australian Centre for Robotic Vision), Kejie Li (The University of Adelaide; Australian Centre for Robotic Vision), Harsh Agarwal (Indian Institute of Technology (BHU)), and Ian M. Reid (The University of Adelaide; Australian Centre for Robotic Vision)

Detect-and-Track: Efficient Pose Estimation in Videos 350
Rohit Girdhar (The Robotics Institute, Carnegie Mellon University), Georgia Gkioxari (Facebook), Lorenzo Torresani (Facebook; Dartmouth College), Manohar Paluri (Facebook), and Du Tran (Facebook)

Supervision-by-Registration: An Unsupervised Approach to Improve the Precision of Facial Landmark Detectors 360
Xuanyi Dong (CAI, University of Technology Sydney), Shoou-I Yu (Facebook Reality Labs, Pittsburgh), Xinshuo Weng (Facebook Reality Labs, Pittsburgh), Shih-En Wei (Facebook Reality Labs, Pittsburgh), Yi Yang (CAI, University of Technology Sydney), and Yaser Sheikh (Facebook Reality Labs, Pittsburgh)

Diversity Regularized Spatiotemporal Attention for Video-Based Person Re-identification 369
Shuang Li (CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong), Slawomir Bak (Argo AI), Peter Carr (Argo AI), and Xiaogang Wang (CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong)

Style Aggregated Network for Facial Landmark Detection 379
Xuanyi Dong (University of Technology Sydney), Yan Yan (University of Technology Sydney), Wanli Ouyang (The University of Sydney), and Yi Yang (University of Technology Sydney)

Learning Deep Models for Face Anti-Spoofing: Binary or Auxiliary Supervision 389
Yaojie Liu (Michigan State University), Amin Jourabloo (Michigan State University), and Xiaoming Liu (Michigan State University)

Deep Cost-Sensitive and Order-Preserving Feature Learning for Cross-Population Age Estimation 399
Kai Li (National Laboratory of Pattern Recognition; University of Chinese Academy of Sciences), Junliang Xing (National Laboratory of Pattern Recognition; University of Chinese Academy of Sciences), Chi Su (KingSoft Ltd.), Weiming Hu (National Laboratory of Pattern Recognition; University of Chinese Academy of Sciences), Junliang Xing (National Laboratory of Pattern Recognition; University of Chinese Academy of Sciences), Yundong Zhang (Vimicro AI Chip Technology Corporation), and Stephen Maybank (Birkbeck University of London)

First-Person Hand Action Benchmark with RGB-D Videos and 3D Hand Pose Annotations 409
Guillermo Garcia-Hernando (Imperial College London), Shanxin Yuan (Imperial College London), Seungryul Baek (Imperial College London), and Tae-Kyun Kim (Imperial College London)

A Pose-Sensitive Embedding for Person Re-identification with Expanded Cross Neighborhood Re-ranking 420
M. Saquib Sarfraz (Karlsruhe Institute of Technology), Arne Schumann (Fraunhofer IOSB Karlsruhe, Germany; Karlsruhe Institute of Technology), Andreas Eberle (Karlsruhe Institute of Technology), and Rainer Stiefelhagen (Karlsruhe Institute of Technology)

Disentangling 3D Pose in a Dendritic CNN for Unconstrained 2D Face Alignment 430
Amit Kumar (University of Maryland-College Park, USA) and Rama Chellappa (University of Maryland-College Park, USA)

A Hierarchical Generative Model for Eye Image Synthesis and Eye Gaze Estimation 440
Kang Wang (Rensselaer Polytechnic Institute, Troy, NY, USA), Rui Zhao (Rensselaer Polytechnic Institute, Troy, NY, USA), and Qiang Ji (Rensselaer Polytechnic Institute, Troy, NY, USA)

MiCT: Mixed 3D/2D Convolutional Tube for Human Action Recognition 449
Yizhou Zhou (University of Science and Technology of China), Xiaoyan Sun (Microsoft Research Asia), Zheng-Jun Zha (University of Science and Technology of China), and Wenjun Zeng (Microsoft Research Asia)

Learning to Estimate 3D Human Pose and Shape from a Single Color Image 459
Georgios Pavlakos (University of Pennsylvania), Luyang Zhu (Peking University), Xiaowei Zhou (Zhejiang University), and Kostas Daniilidis (University of Pennsylvania)

Glimpse Clouds: Human Activity Recognition from Unstructured Feature Points 469
Fabien Baradel (Univ. Lyon, INSA-Lyon, CNRS, LIRIS, France), Christian Wolf (Univ. Lyon, INSA-Lyon, CNRS, LIRIS, F-69621, France;), Julien Mille (Laboratoire d'Informatique de l'Univ. de Tours, INSA Centre Val de Loire, France), and Graham W. Taylor (Univ. of Guelph, Canada; Vector Institute, Canada)

Context-Aware Deep Feature Compression for High-Speed Visual Tracking 479
Jongwon Choi (Seoul National University), Hyung Jin Chang (Imperial College London; University of Birmingham), Tobias Fischer (Imperial College London), Sangdoo Yun (Seoul National University; Clova AI Research, NAVER Corp.), Kyuewang Lee (Seoul National University), Jiyeoup Jeong (Seoul National University), Yiannis Demiris (Imperial College London), and Jin Young Choi (Seoul National University)

Correlation Tracking via Joint Discrimination and Reliability Learning 489
Chong Sun (Dalian University of Technology, China), Dong Wang (Dalian University of Technology, China), Huchuan Lu (Dalian University of Technology, China), and Ming-Hsuan Yang (University of California, Merced, USA)

PhaseNet for Video Frame Interpolation 498
Simone Meyer (ETH Zurich), Simone Meyer (Disney Research), Abdelaziz Djelouah (Disney Research), Brian McWilliams (Disney Research), Alexander Sorkine-Hornung (Disney Research), Markus Gross (ETH Zurich; Disney Research), and Christopher Schroers (Disney Research)

The Best of Both Worlds: Combining CNNs and Geometric Constraints for Hierarchical Motion Segmentation 508
Pia Bideau (University of Massachusetts Amherst), Aruni RoyChowdhury (University of Massachusetts Amherst), Rakesh R. Menon (University of Massachusetts Amherst), and Erik Learned-Miller (University of Massachusetts Amherst)

Hyperparameter Optimization for Tracking with Continuous Deep Q-Learning 518
Xingping Dong (Beijing Institute of Technology, China), Jianbing Shen (Beijing Institute of Technology, China; Inception Institute of Artificial Intelligence, Abu Dhabi, UAE), Wenguan Wang (Beijing Institute of Technology, China), Yu Liu (Beijing Institute of Technology, China), Ling Shao (Inception Institute of Artificial Intelligence, Abu Dhabi, UAE; University of East Anglia, U.K.), and Fatih Porikli (Australian National University, Australia)

Scale-Transferrable Object Detection 528
Peng Zhou (Shanghai Jiao Tong University, China), Bingbing Ni (Shanghai Jiao Tong University, China), Cong Geng (Shanghai Jiao Tong University, China), Jianguo Hu (Minivision, China), and Yi Xu (Shanghai Jiao Tong University, China)

A Prior-Less Method for Multi-face Tracking in Unconstrained Videos 538
Chung-Ching Lin (IBM Research AI) and Ying Hung (Rutgers University)

End-to-End Flow Correlation Tracking with Spatial-Temporal Attention 548
Zheng Zhu (Chinese Academy of Sciences, Beijing, China; University of Chinese Academy of Sciences, Beijing, China), Wei Wu (SenseTime Group Limited, Beijing, China), Wei Zou (Chinese Academy of Sciences, Beijing, China; University of Chinese Academy of Sciences, Beijing, China; TianJin Intelligent Tech.Institute of CASIA Co.,Ltd, Tianjin, China), and Junjie Yan (SenseTime Group Limited, Beijing, China)

Deep Texture Manifold for Ground Terrain Recognition 558
Jia Xue (Rutgers University), Hang Zhang (Rutgers University; Amazon AI), and Kristin Dana (Rutgers University)

Learning Superpixels with Segmentation-Aware Affinity Loss 568
Wei-Chih Tu (National Taiwan University), Ming-Yu Liu (NVIDIA), Varun Jampani (NVIDIA), Deqing Sun (NVIDIA), Shao-Yi Chien (National Taiwan University), Ming-Hsuan Yang (NVIDIA, UC Merced), and Jan Kautz (NVIDIA)

Interactive Image Segmentation with Latent Diversity 577
Zhuwen Li (Intel Labs), Qifeng Chen (Intel Labs), and Vladlen Koltun (Intel Labs)

The Unreasonable Effectiveness of Deep Features as a Perceptual Metric 586
Richard Zhang (UC Berkeley), Phillip Isola (UC Berkeley; OpenAI), Alexei A. Efros (UC Berkeley), Eli Shechtman (Adobe Research), and Oliver Wang (Adobe Research)

Local Descriptors Optimized for Average Precision 596
Kun He (Boston Univesity), Yan Lu (Honda Research Institute USA), and Stan Sclaroff (Boston University)

Recovering Realistic Texture in Image Super-Resolution by Deep Spatial Feature Transform 606
Xintao Wang (CUHK - SenseTime Joint Lab, The Chinese University of Hong Kong), Ke Yu (CUHK - SenseTime Joint Lab, The Chinese University of Hong Kong), Chao Dong (SenseTime Research), and Chen Change Loy (CUHK - SenseTime Joint Lab, The Chinese University of Hong Kong)

Deep Extreme Cut: From Extreme Points to Object Segmentation 616
K.-K. Maninis (ETH Zürich, Switzerland), S. Caelles (ETH Zürich, Switzerland), J. Pont-Tuset (ETH Zürich, Switzerland), and L. Van Gool (ETH Zürich, Switzerland)

Learning to Parse Wireframes in Images of Man-Made Environments 626
Kun Huang (ShanghaiTech University), Yifan Wang (ShanghaiTech University), Zihan Zhou (The Pennsylvania State University), Tianjiao Ding (ShanghaiTech University), Shenghua Gao (ShanghaiTech University), and Yi Ma (University of California, Berkeley)

Occlusion-Aware Rolling Shutter Rectification of 3D Scenes 636
Subeesh Vasu (Indian Institute of Technology Madras), Mahesh Mohan M. R. (Indian Institute of Technology Madras), and A. N. Rajagopalan (Indian Institute of Technology Madras)

Content-Sensitive Supervoxels via Uniform Tessellations on Video Manifolds 646
Ran Yi (Tsinghua University, China), Yong-Jin Liu (Tsinghua University, China), and Yu-Kun Lai (Cardiff University, UK)

Intrinsic Image Transformation via Scale Space Decomposition 656
Lechao Cheng (Zhejiang University), Chengyi Zhang (Zhejiang University), and Zicheng Liao (Zhejiang University; Alibaba - Zhejiang University Joint Institute of Frontier Technologies)

Learned Shape-Tailored Descriptors for Segmentation 666
Naeemullah Khan (King Abdullah University of Science and Technology (KAUST), Saudi Arabia) and Ganesh Sundaramoorthi (King Abdullah University of Science and Technology (KAUST), Saudi Arabia)

PAD-Net: Multi-tasks Guided Prediction-and-Distillation Network for Simultaneous Depth Estimation and Scene Parsing 675
Dan Xu (The University of Trento), Wanli Ouyang (The University of Sydney), Xiaogang Wang (The Chinese University of Hong Kong), and Nicu Sebe (The University of Trento)

Multi-image Semantic Matching by Mining Consistent Features 685
Qianqian Wang (Zhejiang University), Xiaowei Zhou (Zhejiang University), and Kostas Daniilidis (University of Pennsylvania)

Density-Aware Single Image De-raining Using a Multi-stream Dense Network 695
He Zhang (Rutgers University) and Vishal M. Patel (Rutgers University)

Joint Cuts and Matching of Partitions in One Graph 705
Tianshu Yu (Arizona State Univ.), Junchi Yan (Shanghai Jiao Tong University; IBM Research), Jieyi Zhao (Univ. of Texas at Houston), and Baoxin Li (Arizona State Univ.)

Progressive Attention Guided Recurrent Network for Salient Object Detection 714
Xiaoning Zhang (Dalian University of Technology, China), Tiantian Wang (Dalian University of Technology, China), Jinqing Qi (Dalian University of Technology, China), Huchuan Lu (Dalian University of Technology, China), and Gang Wang (Alibaba AILabs, China)

Fast and Accurate Single Image Super-Resolution via Information Distillation Network 723
Zheng Hui (Xidian University, Xi'an, China), Xiumei Wang (Xidian University, Xi'an, China), and Xinbo Gao (Xidian University, Xi'an, China)

Hallucinated-IQA: No-Reference Image Quality Assessment via Adversarial Learning 732
Kwan-Yee Lin (Peking University) and Guanxiang Wang (Peking University)

NAG: Network for Adversary Generation 742
Konda Reddy Mopuri (Indian Institute of Science, Bangalore, India), Utkarsh Ojha (Indian Institute of Science, Bangalore, India), Utsav Garg (Indian Institute of Science, Bangalore, India), and R. Venkatesh Babu (Indian Institute of Science, Bangalore, India)

Dynamic-Structured Semantic Propagation Network 752
Xiaodan Liang (Carnegie Mellon University), Eric Xing (Petuum Inc), and Hongfei Zhou (Wumii Tech. Limited)

Cross-Domain Self-Supervised Multi-task Feature Learning Using Synthetic Imagery 762
Zhongzheng Ren (University of California, Davis) and Yong Jae Lee (University of California, Davis)

A Two-Step Disentanglement Method 772
Naama Hadad (Tel Aviv University), Lior Wolf (Facebook AI Research and Tel Aviv University), and Moni Shahar (Tel Aviv University)

Robust Facial Landmark Detection via a Fully-Convolutional Local-Global Context Network 781
Daniel Merget (Technical University of Munich), Matthias Rock (Technical University of Munich), and Gerhard Rigoll (Technical University of Munich)

Decorrelated Batch Normalization 791
Lei Huangi (Beihang University, P.R.China), Lei Huangi (University of Michigan, Ann Arbor), Dawei Yang (University of Michigan, Ann Arbor), Bo Lang (Beihang University, P.R.China), and Jia Deng (University of Michigan, Ann Arbor)

Learning to Sketch with Shortcut Cycle Consistency 801
Jifei Song (SketchX, Queen Mary University of London), Kaiyue Pang (SketchX, Queen Mary University of London), Yi-Zhe Song (SketchX, Queen Mary University of London), Tao Xiang (SketchX, Queen Mary University of London), and Timothy M. Hospedales (SketchX, Queen Mary University of London; The University of Edinburgh)

Towards a Mathematical Understanding of the Difficulty in Learning with Feedforward Neural Networks 811
Hao Shen (fortiss - The Research Institute of the Free State of Bavaria, Germany)

FaceID-GAN: Learning a Symmetry Three-Player GAN for Identity-Preserving Face Synthesis 821
Yujun Shen (CUHK - SenseTime Joint Lab, The Chinese University of Hong Kong), Ping Luo (CUHK - SenseTime Joint Lab, The Chinese University of Hong Kong), Ping Luo (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences), Junjie Yan (SenseTime Research), Xiaogang Wang (CUHK - SenseTime Joint Lab, The Chinese University of Hong Kong), and Xiaoou Tang (CUHK - SenseTime Joint Lab, The Chinese University of Hong Kong)

A Constrained Deep Neural Network for Ordinal Regression 831
Chi Keong Goh (Rolls-Royce Advanced Technology, Singapore), Yanzhu Liu (Nanyang Technological University, Singapore), and Adams Wai Kin Kong (Nanyang Technological University, Singapore)

Modulated Convolutional Networks 840
Xiaodi Wang (Beihang University, Beijing, China), Baochang Zhang (Beihang University, Beijing, China), Ce Li (China University of Mining & Technology, Beijing, China), Rongrong Ji (Xiamen University, Fujian, China), Jungong Han (Lancaster University, UK), Jianzhuang Liu (Huawei Technologies Co. Ltd., China), and Xianbin Cao (Beihang University, Beijing, China)

Learning Steerable Filters for Rotation Equivariant CNNs 849
Maurice Weiler (AMLab / QUVA Lab, University of Amsterdam), Maurice Weiler (HCI/IWR, University of Heidelberg), Fred A. Hamprecht (HCI/IWR, University of Heidelberg), and Martin Storath (HCI/IWR, University of Heidelberg)

Efficient Interactive Annotation of Segmentation Datasets with Polygon-RNN++ 859
David Acuna (University of Toronto; NVIDIA), Huan Ling (University of Toronto; Vector Institute), Amlan Kar (University of Toronto; Vector Institute), and Sanja Fidler (University of Toronto; Vector Institute)

SplineCNN: Fast Geometric Deep Learning with Continuous B-Spline Kernels 869
Matthias Fey (TU Dortmund University), Jan Eric Lenssen (TU Dortmund University), Frank Weichert (TU Dortmund University), and Heinrich Müller (TU Dortmund University)

GAGAN: Geometry-Aware Generative Adversarial Networks 878
Jean Kossaifi (Imperial College London), Linh Tran (Imperial College London), Yannis Panagakis (Imperial College London; Middlesex University London), and Maja Pantic (Imperial College London)

On the Robustness of Semantic Segmentation Models to Adversarial Attacks 888
Anurag Arnab (University of Oxford), Ondrej Miksik (University of Oxford; Emotech Labs), and Philip H.S. Torr (University of Oxford)

Feedback-Prop: Convolutional Neural Network Inference Under Partial Evidence 898
Tianlu Wang (University of Virginia), Kota Yamaguchi (CyberAgent, Inc.), and Vicente Ordonez (University of Virginia)

Super-Resolving Very Low-Resolution Face Images with Supplementary Attributes 908
Xin Yu (Australian National University), Basura Fernando (Australian National University), Richard Hartley (Australian National University), and Fatih Porikli (Australian National University)

Frustum PointNets for 3D Object Detection from RGB-D Data 918
Charles R. Qi (Stanford University), Wei Liu (Nuro, Inc.), Chenxia Wu (Nuro, Inc.), Hao Su (UC San Diego), and Leonidas J. Guibas (Stanford University)

W2F: A Weakly-Supervised to Fully-Supervised Framework for Object Detection 928
Yongqiang Zhang (King Abdullah University of Science and Technology (KAUST)), Yongqiang Zhang (Harbin Institute of Technology (HIT)), Yancheng Bai (King Abdullah University of Science and Technology (KAUST); Institute of Software, Chinese Academy of Sciences (CAS)), Mingli Ding (Harbin Institute of Technology (HIT)), Yongqiang Li (Harbin Institute of Technology (HIT)), and Bernard Ghanem (King Abdullah University of Science and Technology (KAUST))

3D Object Detection with Latent Support Surfaces 937
Zhile Ren (Brown University) and Erik B. Sudderth (University of California, Irvine)

Towards Faster Training of Global Covariance Pooling Networks by Iterative Matrix Square Root Normalization 947
Peihua Li (Dalian University of Technology), Jiangtao Xie (Dalian University of Technology), Qilong Wang (Dalian University of Technology), and Zilin Gao (Dalian University of Technology)

Recurrent Scene Parsing with Perspective Understanding in the Loop 956
Shu Kong (University of California, Irvine) and Charless Fowlkes (University of California, Irvine)

Improving Occlusion and Hard Negative Handling for Single-Stage Pedestrian Detectors 966
Junhyug Noh (Seoul National University, Seoul, Korea), Soochan Lee (Seoul National University, Seoul, Korea), Beomsu Kim (Seoul National University, Seoul, Korea), and Gunhee Kim (Seoul National University, Seoul, Korea)

Learning to Act Properly: Predicting and Explaining Affordances from Images 975
Ching-Yao Chuang (University of Toronto), Jiaman Li (University of Toronto), Antonio Torralba (Massachusetts Institute of Technology), and Sanja Fidler (University of Toronto)

Pointwise Convolutional Neural Networks 984
Binh-Son Hua (The University of Tokyo), Minh-Khoi Tran (Singapore University of Technology and Design), and Sai-Kit Yeung (Singapore University of Technology and Design)

Image-Image Domain Adaptation with Preserved Self-Similarity and Domain-Dissimilarity for Person Re-identification 994
Weijian Deng (University of Chinese Academy of Sciences), Liang Zheng (University of Chinese Academy of Sciences; Singapore University of Technology and Design), Qixiang Ye (University of Chinese Academy of Sciences), Guoliang Kang (University of Technology Sydney), Yi Yang (University of Technology Sydney), and Jianbin Jiao (University of Chinese Academy of Sciences)

A Generative Adversarial Approach for Zero-Shot Learning from Noisy Texts 1004
Yizhe Zhu (Rutgers University), Mohamed Elhoseiny (Facebook AI Research), Bingchen Liu (Rutgers University), Xi Peng (Rutgers University), and Ahmed Elgammal (Rutgers University)

Tensorize, Factorize and Regularize: Robust Visual Relationship Learning 1014
Seong Jae Hwang (University of Wisconsin - Madison, Madison, WI), Hyunwoo J. Kim (University of Wisconsin - Madison, Madison, WI), Sathya N. Ravi (University of Wisconsin - Madison, Madison, WI), Maxwell D. Collins (University of Wisconsin - Madison, Madison, WI), Zirui Tao (University of Wisconsin - Madison, Madison, WI), and Vikas Singh (University of Wisconsin - Madison, Madison, WI)

Transductive Unbiased Embedding for Zero-Shot Learning 1024
Jie Song (Zhejiang University, Hangzhou, China), Chengchao Shen (Zhejiang University, Hangzhou, China), Yezhou Yang (Arizona State University, Tempe, USA), Yang Liu (Alibaba Group, Hangzhou, China), and Mingli Song (Zhejiang University, Hangzhou, China)

Hierarchical Novelty Detection for Visual Object Recognition 1034
Kibok Lee (University of Michigan, Ann Arbor, MI, USA), Kimin Lee (Korea Advanced Institute of Science and Technology, Daejeon, Korea), Kyle Min (University of Michigan, Ann Arbor, MI, USA), Yuting Zhang (University of Michigan, Ann Arbor, MI, USA), Jinwoo Shin (Korea Advanced Institute of Science and Technology, Daejeon, Korea), and Honglak Lee (University of Michigan, Ann Arbor, MI, USA; Google Brain, Mountain View, CA, USA)

Zero-Shot Visual Recognition Using Semantics-Preserving Adversarial Embedding Networks 1043
Long Chen (Zhejiang University), Hanwang Zhang (Nanyang Technological University; Tencent AI Lab), Jun Xiao (Zhejiang University), Wei Liu (Tencent AI Lab), and Shih-Fu Chang (Columbia University)

Learning Rich Features for Image Manipulation Detection 1053
Peng Zhou (University of Maryland, College Park), Xintong Han (University of Maryland, College Park), Vlad I. Morariu (Adobe Research), and Larry S. Davis (University of Maryland, College Park)

Human Semantic Parsing for Person Re-identification 1062
Mahdi M. Kalayeh (University of Central Florida), Emrah Basaran (Istanbul Technical University), Muhittin Gökmen (MEF University), Mustafa E. Kamasak (Istanbul Technical University), and Mubarak Shah (University of Central Florida)

Stacked Latent Attention for Multimodal Reasoning 1072
Haoqi Fan (Facebook Research) and Jiatong Zhou (Facebook Research)

R-FCN-3000 at 30fps: Decoupling Detection and Classification 1081
Bharat Singh (University of Maryland, College Park), Hengduo Li (Fudan University), Abhishek Sharma (Gobasco AI Labs), and Larry S. Davis (University of Maryland, College Park)

CSRNet: Dilated Convolutional Neural Networks for Understanding the Highly Congested Scenes 1091
Yuhong Li (University of Illinois at Urbana-Champaign), Yuhong Li (Beijing University of Posts and Telecommunications), Xiaofan Zhang (University of Illinois at Urbana-Champaign), and Deming Chen (University of Illinois at Urbana-Champaign)

Revisiting Knowledge Transfer for Training Object Class Detectors 1101
Jasper R. R. Uijlings (Google AI Perception), S. Popov (Google AI Perception), and V. Ferrari (Google AI Perception)

Deep Sparse Coding for Invariant Multimodal Halle Berry Neurons 1111
Edward Kim (Villanova University, PA; Los Alamos National Laboratory, Los Alamos, NM), Darryl Hannan (Villanova University, PA; Los Alamos National Laboratory, Los Alamos, NM), and Garrett Kenyon (Los Alamos National Laboratory, Los Alamos, NM)

On the Convergence of PatchMatch and Its Variants 1121
Thibaud Ehret (Université Paris-Saclay, France) and Pablo Arias (CNRS, Université Paris-Saclay, France)

Rethinking the Faster R-CNN Architecture for Temporal Action Localization 1130
Yu-Wei Chao (University of Michigan, Ann Arbor), Sudheendra Vijayanarasimhan (Google Research), Bryan Seybold (Google Research), David A. Ross (Google Research), Jia Deng (University of Michigan, Ann Arbor), and Rahul Sukthankar (Google Research)

MoNet: Deep Motion Exploitation for Video Object Segmentation 1140
Huaxin Xiao (National University of Defense Technology), Huaxin Xiao (National University of Singapore), Jiashi Feng (National University of Singapore), Guosheng Lin (Nanyang Technological University), Yu Liu (National University of Defense Technology), and Maojun Zhang (National University of Defense Technology)

Video Representation Learning Using Discriminative Pooling 1149
Jue Wang (Data61/CSIRO), Jue Wang (The Australian National University, Canberra, Australia), Anoop Cherian (Australian Centre for Robotic Vision; The Australian National University, Canberra, Australia; MERL Cambridge, MA), Fatih Porikli (Australian Centre for Robotic Vision; The Australian National University, Canberra, Australia), and Stephen Gould (Australian Centre for Robotic Vision; The Australian National University, Canberra, Australia)

Recognizing Human Actions as the Evolution of Pose Estimation Maps 1159
Mengyuan Liu (Nanyang Technological University, Singapore) and Junsong Yuan (SUNY at Buffalo, Buffalo NY)

Video Person Re-identification with Competitive Snippet-Similarity Aggregation and Co-attentive Snippet Embedding 1169
Dapeng Chen (The Chinese University of Hong Kong; Xi'an Jiaotong University), Hongsheng Li (The Chinese University of Hong Kong), Tong Xiao (The Chinese University of Hong Kong), Shuai Yi (SenseTime Research), and Xiaogang Wang (The Chinese University of Hong Kong)

Mask-Guided Contrastive Attention Model for Person Re-identification 1179
Chunfeng Song (National Laboratory of Pattern Recognition (NLPR); University of Chinese Academy of Sciences (UCAS)), Yan Huang (National Laboratory of Pattern Recognition (NLPR); University of Chinese Academy of Sciences (UCAS)), Wanli Ouyang (University of Sydney), and Liang Wang (National Laboratory of Pattern Recognition (NLPR); Institute of Automation, Chinese Academy of Sciences (CASIA); University of Chinese Academy of Sciences (UCAS))

Blazingly Fast Video Object Segmentation with Pixel-Wise Metric Learning 1189
Yuhua Chen (Computer Vision Lab, ETH Zurich), Jordi Pont-Tuset (Computer Vision Lab, ETH Zurich), Alberto Montes (Computer Vision Lab, ETH Zurich), and Luc Van Gool (Computer Vision Lab, ETH Zurich; KU Leuven)

Poster Session P1-2

Learning to Compare: Relation Network for Few-Shot Learning 1199
Flood Sung (Queen Mary University of London), Yongxin Yang (The University of Edinburgh), Li Zhang (University of Oxford), Tao Xiang (Queen Mary University of London), Philip H.S. Torr (University of Oxford), and Timothy M. Hospedales (The University of Edinburgh)

COCO-Stuff: Thing and Stuff Classes in Context 1209
Holger Caesar (University of Edinburgh), Jasper Uijlings (Google AI Perception), and Vittorio Ferrari (Google AI Perception)

Image Generation from Scene Graphs 1219
Justin Johnson (Stanford University; Google Cloud AI), Agrim Gupta (Stanford University), and Li Fei-Fei (Stanford University; Google Cloud AI)

Deep Cauchy Hashing for Hamming Space Retrieval 1229
Yue Cao (Tsinghua University, China), Mingsheng Long (Tsinghua University, China), Bin Liu (Tsinghua University, China), and Jianmin Wang (Tsinghua University, China)

Learning to Look Around: Intelligently Exploring Unseen Environments for Unknown Tasks 1238
Dinesh Jayaraman (UC Berkeley) and Kristen Grauman (UT Austin)

Multi-scale Location-Aware Kernel Representation for Object Detection 1248
Hao Wang (Harbin Institute of Technology, Harbin, China), Qilong Wang (Dalian University of Technology, Dalian, China), Mingqi Gao (Harbin Institute of Technology, Harbin, China), Peihua Li (Dalian University of Technology, Dalian, China), and Wangmeng Zuo (Harbin Institute of Technology, Harbin, China)

Clinical Skin Lesion Diagnosis Using Representations Inspired by Dermatologist Criteria 1258
Jufeng Yang (Nankai University, Tianjin, China), Xiaoxiao Sun (Nankai University, Tianjin, China), Jie Liang (Nankai University, Tianjin, China), and Paul L. Rosin (Cardiff University, Cardiff, UK)

Compare and Contrast: Learning Prominent Visual Differences 1267
Steven Chen (The University of Texas at Austin) and Kristen Grauman (The University of Texas at Austin)

Multi-evidence Filtering and Fusion for Multi-label Classification, Object Detection and Semantic Segmentation Based on Weakly Supervised Learning 1277
Weifeng Ge (The University of Hong Kong), Sibei Yang (The University of Hong Kong), and Yizhou Yu (The University of Hong Kong)

HashGAN: Deep Learning to Hash with Pair Conditional Wasserstein GAN 1287
Yue Cao (Tsinghua University, China), Bin Liu (Tsinghua University, China), Mingsheng Long (Tsinghua University, China), and Jianmin Wang (Tsinghua University, China)

Min-Entropy Latent Model for Weakly Supervised Object Detection 1297
Fang Wan (University of Chinese Academy of Sciences, Beijing), Pengxu Wei (University of Chinese Academy of Sciences, Beijing), Jianbin Jiao (University of Chinese Academy of Sciences, Beijing), Zhenjun Han (University of Chinese Academy of Sciences, Beijing), and Qixiang Ye (University of Chinese Academy of Sciences, Beijing)

MAttNet: Modular Attention Network for Referring Expression Comprehension 1307
Licheng Yu (University of North Carolina at Chapel Hill), Zhe Lin (Adobe Research), Xiaohui Shen (Adobe Research), Jimei Yang (Adobe Research), Xin Lu (Adobe Research), Mohit Bansal (University of North Carolina at Chapel Hill), and Tamara L. Berg (University of North Carolina at Chapel Hill)

AttnGAN: Fine-Grained Text to Image Generation with Attentional Generative Adversarial Networks 1316
Tao Xu (Lehigh University), Pengchuan Zhang (Microsoft Research), Qiuyuan Huang (Microsoft Research), Han Zhang (Rutgers University), Zhe Gan (Duke University), Xiaolei Huang (Lehigh University), and Xiaodong He (JD AI Research)

Adversarial Complementary Learning for Weakly Supervised Object Localization 1325
Xiaolin Zhang (CAI, University of Technology Sydney), Yunchao Wei (University of Illinois Urbana-Champaign), Jiashi Feng (National University of Singapore), Yi Yang (CAI, University of Technology Sydney), and Thomas Huang (University of Illinois Urbana-Champaign)

Conditional Generative Adversarial Network for Structured Domain Adaptation 1335
Weixiang Hong (Nanyang Technological University), Zhenzhen Wang (Nanyang Technological University), Ming Yang (Horizon Robotics, Inc.), and Junsong Yuan (State University of New York at Buffalo)

GroupCap: Group-Based Image Captioning with Structured Relevance and Diversity Constraints 1345
Fuhai Chen (Xiamen University, China), Rongrong Ji (Xiamen University, China), Xiaoshuai Sun (Harbin Institute of Technology, China), Yongjian Wu (Tencent YouTu Lab), and Jinsong Su (Xiamen University, China)

Weakly-Supervised Semantic Segmentation by Iteratively Mining Common Object Features 1354
Xiang Wang (Tsinghua University, China), Shaodi You (DATA61-CSIRO), Shaodi You (Australian National University, Australia), Xi Li (Tsinghua University, China), and Huimin Ma (Tsinghua University, China)

Bootstrapping the Performance of Webly Supervised Semantic Segmentation 1363
Tong Shen (The University of Adelaide, Australia), Guosheng Lin (Nanyang Technological University, Singapore), Chunhua Shen (The University of Adelaide, Australia), and Ian Reid (The University of Adelaide, Australia)

DeepVoting: A Robust and Explainable Deep Network for Semantic Part Detection Under Partial Occlusion 1372
Zhishuai Zhang (Johns Hopkins University), Cihang Xie (Johns Hopkins University), Jianyu Wang (Baidu Research USA), Lingxi Xie (Johns Hopkins University), and Alan L. Yuille (Johns Hopkins University)

Geometry-Aware Scene Text Detection with Instance Transformation Network 1381
Fangfang Wang (Zhejiang University, China), Liming Zhao (Zhejiang University, China), Xi Li (Zhejiang University, China), Xinchao Wang (Stevens Institute of Technology, USA), and Dacheng Tao (University of Sydney, Australia)

Optical Flow Guided Feature: A Fast and Robust Motion Representation for Video Action Recognition 1390
Shuyang Sun (The University of Sydney), Shuyang Sun (SenseTime Research), Zhanghui Kuang (SenseTime Research), Lu Sheng (The Chinese University of Hong Kong), Wanli Ouyang (The University of Sydney), and Wei Zhang (SenseTime Research)

Motion-Guided Cascaded Refinement Network for Video Object Segmentation 1400
Ping Hu (Nanyang Technological University), Gang Wang (Alibaba AI Labs), Xiangfei Kong (Nanyang Technological University), Jason Kuen (Nanyang Technological University), and Yap-Peng Tan (Nanyang Technological University)

A Memory Network Approach for Story-Based Temporal Summarization of 360° Videos 1410
Sangho Lee (Seoul National University), Jinyoung Sung (Seoul National University), Youngjae Yu (Seoul National University), and Gunhee Kim (Seoul National University)

Cube Padding for Weakly-Supervised Saliency Prediction in 360° Videos 1420
Hsien-Tzu Cheng (National Tsing Hua University), Chun-Hung Chao (National Tsing Hua University), Jin-Dong Dong (National Tsing Hua University), Hao-Kai Wen (Taiwan AI Labs), Tyng-Luh Liu (Academia Sinica), and Min Sun (National Tsing Hua University)

Appearance-and-Relation Networks for Video Classification 1430
Limin Wang (Nanjing University, China), Limin Wang (ETH Zurich, Switzerland), Wei Li (Google Research; ETH Zurich, Switzerland), and Luc Van Gool (ETH Zurich, Switzerland)

Excitation Backprop for RNNs 1440
Sarah Adel Bargal (Boston University), Andrea Zunino (Istituto Italiano di Tecnologia), Donghyun Kim (Boston University), Jianming Zhang (Adobe Research), Vittorio Murino (Istituto Italiano di Tecnologia; Universitá di Verona), and Stan Sclaroff (Boston University)

One-Shot Action Localization by Learning Sequence Matching Network 1450
Hongtao Yang (The Australian National University), Xuming He (ShanghaiTech University), and Fatih Porikli (The Australian National University)

Structure Preserving Video Prediction 1460
Jingwei Xu (Shanghai Jiao Tong University, Shanghai, China), Bingbing Ni (Shanghai Jiao Tong University, Shanghai, China), Zefan Li (Shanghai Jiao Tong University, Shanghai, China), Shuo Cheng (Shanghai Jiao Tong University, Shanghai, China), and Xiaokang Yang (Shanghai Jiao Tong University, Shanghai, China)

Person Re-identification with Cascaded Pairwise Convolutions 1470
Yicheng Wang (Wuhan University, China), Zhenzhong Chen (Wuhan University, China), Feng Wu (University of Science and Technology of China, China), and Gang Wang (Alibaba Group, China)

On the Importance of Label Quality for Semantic Segmentation 1479
Aleksandar Zlateski (Massachusetts Institute of Technology), Ronnachai Jaroensri (Massachusetts Institute of Technology), Prafull Sharma (Massachusetts Institute of Technology), and Fredo Durand (Massachusetts Institute of Technology)

Scalable and Effective Deep CCA via Soft Decorrelation 1488
Xiaobin Chang (Queen Mary University of London), Tao Xiang (Queen Mary University of London), and Timothy M. Hospedales (The University of Edinburgh)

Duplex Generative Adversarial Network for Unsupervised Domain Adaptation 1498
Lanqing Hu (Institute of Computing Technology, CAS, Beijing, China; University of Chinese Academy of Sciences, Beijing, China), Meina Kan (Institute of Computing Technology, CAS, Beijing, China), Shiguang Shan (Institute of Computing Technology, CAS, Beijing, China), and Xilin Chen (Institute of Computing Technology, CAS, Beijing, China)

Edit Probability for Scene Text Recognition 1508
Fan Bai (Fudan University, Shanghai, China), Zhanzhan Cheng (Hikvision Research Institute, China), Yi Niu (Hikvision Research Institute, China), Shiliang Pu (Hikvision Research Institute, China), and Shuigeng Zhou (Fudan University, Shanghai, China)

Global Versus Localized Generative Adversarial Nets 1517
Guo-Jun Qi (University of Central Florida and Alibaba Group), Liheng Zhang (University of Central Florida and Alibaba Group), Hao Hu (University of Central Florida and Alibaba Group), Marzieh Edraki (University of Central Florida and Alibaba Group), Jingdong Wang (University of Central Florida and Alibaba Group), and Xian-Sheng Hua (University of Central Florida and Alibaba Group)

MoCoGAN: Decomposing Motion and Content for Video Generation 1526
Sergey Tulyakov (Snap Research NVIDIA), Ming-Yu Liu (Snap Research NVIDIA), Xiaodong Yang (Snap Research NVIDIA), and Jan Kautz (Snap Research NVIDIA)

Recurrent Residual Module for Fast Inference in Videos 1536
Bowen Pan (Shanghai Jiao Tong University), Wuwei Lin (Shanghai Jiao Tong University), Xiaolin Fang (Zhejiang University), Chaoqin Huang (Shanghai Jiao Tong University), Bolei Zhou (Massachusetts Institute of Technology), and Cewu Lu (Shanghai Jiao Tong University)

Improving Landmark Localization with Semi-Supervised Learning 1546
Sina Honari (MILA-University of Montreal), Pavlo Molchanov (NVIDIA), Stephen Tyree (NVIDIA), Pascal Vincent (MILA-University of Montreal; CIFAR; Facebook AI Research), Christopher Pal (MILA-University of Montreal; Ecole Polytechnique of Montreal), and Jan Kautz (NVIDIA)

Adversarial Data Programming: Using GANs to Relax the Bottleneck of Curated Labeled Data 1556
Arghya Pal (Indian Institute of Technology, Hyderabad, India) and Vineeth N Balasubramanian (Indian Institute of Technology, Hyderabad, India)

Stochastic Variational Inference with Gradient Linearization 1566
Tobias Plötz (TU Darmstadt), Anne S. Wannenwetsch (TU Darmstadt), and Stefan Roth (TU Darmstadt)

Multi-label Zero-Shot Learning with Structured Knowledge Graphs 1576
Chung-Wei Lee (National Taiwan University, Taipei, Taiwan), Wei Fang (National Taiwan University, Taipei, Taiwan), Chih-Kuan Yeh (Carnegie Mellon University, Pittsburgh, USA), and Yu-Chiang Frank Wang (National Taiwan University, Taipei, Taiwan)

MorphNet: Fast & Simple Resource-Constrained Structure Learning of Deep Networks 1586
Ariel Gordon (Google AI), Elad Eban (Google AI), Ofir Nachum (Google Brain), Bo Chen (Google AI), Hao Wu (Google AI), Tien-Ju Yang (Google AI; MIT), and Edward Choi (Google AI; Georgia Institute of Technology)

Deep Adversarial Subspace Clustering 1596
Pan Zhou (National University of Singapore, Singapore), Yunqing Hou (National University of Singapore, Singapore), and Jiashi Feng (National University of Singapore, Singapore)

Towards Human-Machine Cooperation: Self-Supervised Sample Mining for Object Detection 1605
Keze Wang (Sun Yat-sen University), Keze Wang (The Hong Kong Polytechnic University), Xiaopeng Yan (Sun Yat-sen University), Dongyu Zhang (Sun Yat-sen University), Lei Zhang (The Hong Kong Polytechnic University), Liang Lin (Sun Yat-sen University), and Liang Lin (SenseTime Research)

Discrete-Continuous ADMM for Transductive Inference in Higher-Order MRFs 1614
Emanuel Laude (Technical University of Munich), Jan-Hendrik Lange (Max Planck Institute for Informatics, Saarbrücken), Jonas Schüpfer (Technical University of Munich), Csaba Domokos (Technical University of Munich), Laura Leal-Taixé (Technical University of Munich), Frank R. Schmidt (Technical University of Munich; Bosch Center for Artificial Intelligence), Bjoern Andres (Max Planck Institute for Informatics, Saarbrücken; Bosch Center for Artificial Intelligence; University of Tübingen), and Daniel Cremers (Technical University of Munich)

Robust Physical-World Attacks on Deep Learning Visual Classification 1625
Kevin Eykholt (University of Michigan, Ann Arbor), Ivan Evtimov (University of Washington), Earlence Fernandes (University of Washington), Bo Li (University of California, Berkeley), Amir Rahmati (Samsung Research America and Stony Brook University), Chaowei Xiao (University of Michigan, Ann Arbor), Atul Prakash (University of Michigan, Ann Arbor), Tadayoshi Kohno (University of Washington), and Dawn Song (University of California, Berkeley)

Generating a Fusion Image: One's Identity and Another's Shape 1635
Donggyu Joo (KAIST, South Korea), Doyeon Kim (KAIST, South Korea), and Junmo Kim (KAIST, South Korea)

Learning to Promote Saliency Detectors 1644
Yu Zeng (Dalian University of Technology, China), Huchuan Lu (Dalian University of Technology, China), Lihe Zhang (Dalian University of Technology, China), Mengyang Feng (Dalian University of Technology, China), and Ali Borji (University of Central Florida, USA)

Image Super-Resolution via Dual-State Recurrent Networks 1654
Wei Han (University of Illinois at Urbana-Champaign), Shiyu Chang (IBM Research), Ding Liu (University of Illinois at Urbana-Champaign), Mo Yu (IBM Research), Michael Witbrock (IBM Research), and Thomas S. Huang (University of Illinois at Urbana-Champaign)

Deep Back-Projection Networks for Super-Resolution 1664
Muhammad Haris (Toyota Technological Institute, Japan), Greg Shakhnarovich (Toyota Technological Institute at Chicago, USA), and Norimichi Ukita (Toyota Technological Institute, Japan)

Focus Manipulation Detection via Photometric Histogram Analysis 1674
Scott McCloskey (Honeywell ACS Labs Golden Valley, USA), Can Chen (University of Delaware Newark, USA), and Jingyi Yu (ShanghaiTech University Shanghai, China)

Compassionately Conservative Balanced Cuts for Image Segmentation 1683
Nathan D. Cahill (Rochester Institute of Technology), Tyler L. Hayes (Rochester Institute of Technology), Renee T. Meinhold (Rochester Institute of Technology), and John F. Hamilton (Rochester Institute of Technology)

A High-Quality Denoising Dataset for Smartphone Cameras 1692
Abdelrahman Abdelhamed (York University), Stephen Lin (Microsoft Research), and Michael S. Brown (York University)

Context-Aware Synthesis for Video Frame Interpolation 1701
Simon Niklaus (Portland State University) and Feng Liu (Portland State University)

Salient Object Detection Driven by Fixation Prediction 1711
Wenguan Wang (Beijing Institute of Technology, China), Jianbing Shen (Beijing Institute of Technology, China; Inception Institute of Artificial Intelligence, Abu Dhabi, UAE), Xingping Dong (Beijing Institute of Technology, China), and Ali Borji (University of Central Florida, USA)

Enhancing the Spatial Resolution of Stereo Images Using a Parallax Prior 1721
Daniel S. Jeon (Korea Advanced Institute of Science and Technology (KAIST)), Seung-Hwan Baek (Korea Advanced Institute of Science and Technology (KAIST)), Inchang Choi (Korea Advanced Institute of Science and Technology (KAIST)), and Min H. Kim (Korea Advanced Institute of Science and Technology (KAIST))

HATS: Histograms of Averaged Time Surfaces for Robust Event-Based Object Classification 1731
Amos Sironi (PROPHESEE, Paris, France), Manuele Brambilla (PROPHESEE, Paris, France), Nicolas Bourdis (PROPHESEE, Paris, France), Xavier Lagorce (PROPHESEE, Paris, France), Ryad Benosman (PROPHESEE, Paris, France), Ryad Benosman (Institut de la Vision, UPMC, Paris, France), and Ryad Benosman (University of Pittsburgh Medical Center / Carnegie Mellon University)

A Bi-Directional Message Passing Model for Salient Object Detection 1741
Lu Zhang (Dalian University of Technology, China), Ju Dai (Dalian University of Technology, China), Huchuan Lu (Dalian University of Technology, China), You He (Naval Aviation University, China), and Gang Wang (Alibaba AILabs, China)

Matching Pixels Using Co-occurrence Statistics 1751
Rotal Kat (Tel-Aviv University, Israel), Roy Jevnisek (Tel-Aviv University, Israel), and Shai Avidan (Tel-Aviv University, Israel)

SeedNet: Automatic Seed Generation with Deep Reinforcement Learning for Robust Interactive Segmentation 1760
Kyoung Mu Lee (Seoul National University, Korea), Heesoo Myeong (Seoul National University, Korea), and Gwangmo Song (Seoul National University, Korea)

Jerk-Aware Video Acceleration Magnification 1769
Shoichiro Takeda (NTT Media Intelligence Laboratories, Japan), Kazuki Okami (NTT Media Intelligence Laboratories, Japan), Dan Mikami (NTT Media Intelligence Laboratories, Japan), Megumi Isogai (NTT Media Intelligence Laboratories, Japan), and Hideaki Kimata (NTT Media Intelligence Laboratories, Japan)

Defense Against Adversarial Attacks Using High-Level Representation Guided Denoiser 1778
Fangzhou Liao (Tsinghua University, China), Ming Liang (Tsinghua University, China), Yinpeng Dong (Tsinghua University, China), Tianyu Pang (Tsinghua University, China), Xiaolin Hu (Tsinghua University, China), and Jun Zhu (Tsinghua University, China)

Stacked Conditional Generative Adversarial Networks for Jointly Learning Shadow Detection and Shadow Removal 1788
Jifeng Wang (Nanjing University of Science and Technology, China), Xiang Li (Nanjing University of Science and Technology, China), and Jian Yang (Nanjing University of Science and Technology, China)

Image Correction via Deep Reciprocating HDR Transformation 1798
Xin Yang (Dalian University of Technology, China; City University of Hong Kong), Ke Xu (Dalian University of Technology, China; City University of Hong Kong), Yibing Song (Tencent AI Lab), Qiang Zhang (Dalian University of Technology, China), Xiaopeng Wei (Dalian University of Technology, China), and Rynson W.H. Lau (City University of Hong Kong)

PieAPP: Perceptual Image-Error Assessment Through Pairwise Preference 1808
Ekta Prashnani (University of California, Santa Barbara), Hong Cai (University of California, Santa Barbara), Yasamin Mostofi (University of California, Santa Barbara), and Pradeep Sen (University of California, Santa Barbara)

Normalized Cut Loss for Weakly-Supervised CNN Segmentation 1818
Meng Tang (University of Waterloo, Canada), Meng Tang (Disney Research, Zurich, Switzerland), Abdelaziz Djelouah (Disney Research, Zurich, Switzerland), Federico Perazzi (Disney Research, Zurich, Switzerland), Yuri Boykov (University of Waterloo, Canada), and Christopher Schroers (Disney Research, Zurich, Switzerland)

ISTA-Net: Interpretable Optimization-Inspired Deep Network for Image Compressive Sensing 1828
Jian Zhang (King Abdullah University of Science and Technology (KAUST), Saudi Arabia) and Bernard Ghanem (King Abdullah University of Science and Technology (KAUST), Saudi Arabia)

Fast End-to-End Trainable Guided Filter 1838
Huikai Wu (Chinese Academy of Sciences, Beijing, China), Huikai Wu (University of Chinese Academy of Sciences, Beijing, China), Shuai Zheng (eBay Research), Junge Zhang (NLPR, Institute of Automation, Chinese Academy of Sciences, Beijing, China; NLPR, Institute of Automation, Chinese Academy of Sciences, Beijing, China), and Kaiqi Huang (NLPR, Institute of Automation, Chinese Academy of Sciences, Beijing, China; NLPR, Institute of Automation, Chinese Academy of Sciences, Beijing, China)

Disentangling Structure and Aesthetics for Style-Aware Image Completion 1848
Andrew Gilbert (University of Surrey), John Collomosse (University of Surrey; Creative Intelligence Lab, Adobe Research), Hailin Jin (Creative Intelligence Lab, Adobe Research), and Brian Price (Creative Intelligence Lab, Adobe Research)

Learning a Discriminative Feature Network for Semantic Segmentation 1857
Changqian Yu (Huazhong University of Science and Technology), Jingbo Wang (Key Laboratory of Machine Perception, Peking University), Chao Peng (Megvii Inc. (Face++)), Changxin Gao (Huazhong University of Science and Technology), Gang Yu (Megvii Inc. (Face++)), and Nong Sang (Huazhong University of Science and Technology)

Kernelized Subspace Pooling for Deep Local Descriptors 1867
Xing Wei (Xi'an Jiaotong University), Yue Zhang (Xi'an Jiaotong University), Yihong Gong (Xi'an Jiaotong University), and Nanning Zheng (Xi'an Jiaotong University)

pOSE: Pseudo Object Space Error for Initialization-Free Bundle Adjustment 1876
Christopher Zach (Toshiba Research Europe) and Je Hyeong Hong (University of Cambridge)

Deformable Shape Completion with Graph Convolutional Autoencoders 1886
Or Litany (Tel Aviv University; Google Research), Alex Bronstein (Tel Aviv University; Technion), Michael Bronstein (USI Lugano), and Ameesh Makadia (Google Research)

Learning from Millions of 3D Scans for Large-Scale 3D Face Recognition 1896
Syed Zulqarnain Gilani (The University of Western Australia) and Ajmal Mian (The University of Western Australia)

CarFusion: Combining Point Tracking and Part Detection for Dynamic 3D Reconstruction of Vehicles 1906
N Dinesh Reddy (Carnegie Mellon University), Minh Vo (Carnegie Mellon University), and Srinivasa G. Narasimhan (Carnegie Mellon University)

Deep Material-Aware Cross-Spectral Stereo Matching 1916
Tiancheng Zhi (Carnegie Mellon University), Bernardo R. Pires (Carnegie Mellon University), Martial Hebert (Carnegie Mellon University), and Srinivasa G. Narasimhan (Carnegie Mellon University)

Augmenting Crowd-Sourced 3D Reconstructions Using Semantic Detections 1926
True Price (UNC Chapel Hill), Johannes L. Schönberger (ETH Zürich), Zhen Wei (UNC Chapel Hill), Marc Pollefeys (ETH Zürich; Microsoft), and Jan-Michael Frahm (UNC Chapel Hill)

Matryoshka Networks: Predicting 3D Geometry via Nested Shape Layers 1936
Stefan Roth (TU Darmstadt) and Stephan R. Richter (TU Darmstadt; Intel Labs)

Triplet-Center Loss for Multi-view 3D Object Retrieval 1945
Xinwei He (Huazhong University of Science and Technology), Yang Zhou (Huazhong University of Science and Technology), Zhichao Zhou (Huazhong University of Science and Technology), Song Bai (Huazhong University of Science and Technology), and Xiang Bai (Huazhong University of Science and Technology)

Learning 3D Shape Completion from Laser Scan Data with Weak Supervision 1955
David Stutz (MPI for Intelligent Systems and University of Tübingen; Max-Planck Institute for Informatics, Saarbrücken) and Andreas Geiger (MPI for Intelligent Systems and University of Tübingen; ETH Zürich)

End-to-End Learning of Keypoint Detector and Descriptor for Pose Invariant 3D Matching 1965
Georgios Georgakis (George Mason University, Fairfax VA), Srikrishna Karanam (Siemens Corporate Technology, Princeton NJ), Ziyan Wu (Siemens Corporate Technology, Princeton NJ), Jan Ernst (Siemens Corporate Technology, Princeton NJ), and Jana Košecká (George Mason University, Fairfax VA)

ICE-BA: Incremental, Consistent and Efficient Bundle Adjustment for Visual-Inertial SLAM 1974
Haomin Liu (Baidu), Mingyu Chen (Baidu), Guofeng Zhang (Zhejiang University), Hujun Bao (Zhejiang University), and Yingze Bao (Baidu)

GeoNet: Unsupervised Learning of Dense Depth, Optical Flow and Camera Pose 1983
Zhichao Yin (SenseTime Research) and Jianping Shi (SenseTime Research)

Radially-Distorted Conjugate Translations 1993
James Pritts (CTU in Prague), Zuzana Kukelova (CTU in Prague), Viktor Larsson (Lund University), and Ondrej Chum (CTU in Prague)

Deep Ordinal Regression Network for Monocular Depth Estimation 2002
Huan Fu (The University of Sydney, Australia), Mingming Gong (University of Pittsburgh; Carnegie Mellon University), Chaohui Wang (UPEM, Marne-la-Vallée, France), Kayhan Batmanghelich (University of Pittsburgh), and Dacheng Tao (The University of Sydney, Australia)

Analytical Modeling of Vanishing Points and Curves in Catadioptric Cameras 2012
Pedro Miraldo (Instituto Superior Técnico, Lisboa), Srikumar Ramalingam (University of Utah), and Francisco Eiras (University of Oxford)

Learning Depth from Monocular Videos Using Direct Methods 2022
Chaoyang Wang (Carnegie Mellon University), José Miguel Buenaposada (Carnegie Mellon University; Universidad Rey Juan Carlos), Rui Zhu (Carnegie Mellon University), and Simon Lucey (Carnegie Mellon University)

Salience Guided Depth Calibration for Perceptually Optimized Compressive Light Field 3D Display 2031
Shizheng Wang (Institute of Microelectronics of Chinese Academy of Sciences, China), Wenjuan Liao (Institute of Microelectronics of Chinese Academy of Sciences, China), Phil Surman (Nanyang Technological University, Singapore), Zhigang Tu (Nanyang Technological University, Singapore), Yuanjin Zheng (Nanyang Technological University, Singapore), and Junsong Yuan (University at Buffalo, USA)

MegaDepth: Learning Single-View Depth Prediction from Internet Photos 2041
Zhengqi Li (Cornell University) and Noah Snavely (Cornell University)

LayoutNet: Reconstructing the 3D Room Layout from a Single RGB Image 2051
Chuhang Zou (University of Illinois at Urbana-Champaign), Alex Colburn (Zillow Group), Qi Shan (Zillow Group), and Derek Hoiem (University of Illinois at Urbana-Champaign)

CBMV: A Coalesced Bidirectional Matching Volume for Disparity Estimation 2060
Konstantinos Batsos (Stevens Institute of Technology), Changjiang Cai (Stevens Institute of Technology), and Philippos Mordohai (Stevens Institute of Technology)

Zoom and Learn: Generalizing Deep Stereo Matching to Novel Domains 2070
Jiahao Pang (SenseTime Research), Wenxiu Sun (SenseTime Research), Chengxi Yang (SenseTime Research), Jimmy Ren (SenseTime Research), Ruichao Xiao (SenseTime Research), Jin Zeng (SenseTime Research), and Liang Lin (SenseTime Research; Sun Yat-sen University)

Exploring Disentangled Feature Representation Beyond Face Identification 2080
Yu Liu (The Chinese University of Hong Kong), Fangyin Wei (SenseTime Group Limited), Jing Shao (SenseTime Group Limited), Lu Sheng (The Chinese University of Hong Kong), Junjie Yan (SenseTime Group Limited), and Xiaogang Wang (The Chinese University of Hong Kong)

Learning Facial Action Units from Web Images with Scalable Weakly Supervised Clustering 2090
Kaili Zhao (Beijing University of Posts and Telecom), Wen-Sheng Chu (Carnegie Mellon University), and Aleix M. Martinez (The Ohio State University)

Human Pose Estimation with Parsing Induced Learner 2100
Xuecheng Nie (National University of Singapore, Singapore), Jiashi Feng (National University of Singapore, Singapore), Yiming Zuo (Tsinghua University, Beijing, China), and Shuicheng Yan (National University of Singapore, Singapore; Qihoo 360 AI Institute, Beijing, China)

Multi-level Factorisation Net for Person Re-identification 2109
Xiaobin Chang (Queen Mary University of London), Timothy M. Hospedales (The University of Edinburgh), and Tao Xiang (Queen Mary University of London)

Attention-Aware Compositional Network for Person Re-identification 2119
Jing Xu (SenseNets Technology Limited), Rui Zhao (SenseTime Group Limited), Feng Zhu (SenseNets Technology Limited), Huaming Wang (SenseNets Technology Limited), and Wanli Ouyang (The University of Sydney)

Look at Boundary: A Boundary-Aware Face Alignment Algorithm 2129
Wenyan (Wayne) Wu (Tsinghua University, China), Wenyan (Wayne) Wu (SenseTime Research), Chen Qian (SenseTime Research), Shuo Yang (Amazon Rekognition), Quan Wang (SenseTime Research), Yici Cai (Tsinghua University, China), and Qiang Zhou (Tsinghua University, China)

Demo2Vec: Reasoning Object Affordances from Online Videos 2139
Kuan Fang (Stanford University), Te-Lin Wu (University of Southern California), Daniel Yang (University of Southern California), Silvio Savarese (Stanford University), and Joseph J. Lim (University of Southern California)

Monocular 3D Pose and Shape Estimation of Multiple People in Natural Scenes: The Importance of Multiple Scene Constraints 2148
Andrei Zanfir (Institute of Mathematics of the Romanian Academy), Elisabeta Marinoiu (Institute of Mathematics of the Romanian Academy), and Cristian Sminchisescu (Lund University; Institute of Mathematics of the Romanian Academy)

3D Human Sensing, Action and Emotion Recognition in Robot Assisted Therapy of Children with Autism 2158
Elisabeta Marinoiu (Institute of Mathematics of the Romanian Academy), Mihai Zanfir (Institute of Mathematics of the Romanian Academy), Vlad Olaru (Institute of Mathematics of the Romanian Academy), and Cristian Sminchisescu (Lund University; Institute of Mathematics of the Romanian Academy)

Facial Expression Recognition by De-expression Residue Learning 2168
Huiyuan Yang (State University of New York at Binghamton, USA), Umur Ciftci (State University of New York at Binghamton, USA), and Lijun Yin (State University of New York at Binghamton, USA)

A Causal And-Or Graph Model for Visibility Fluent Reasoning in Tracking Interacting Objects 2178
Yuanlu Xu (University of California, Los Angeles), Lei Qin (Inst. Computing Technology, Chinese Academy of Sciences), Xiaobai Liu (San Diego State University), Jianwen Xie (Hikvision Research Institute, USA), and Song-Chun Zhu (University of California, Los Angeles)

Weakly Supervised Facial Action Unit Recognition Through Adversarial Training 2188
Guozhu Peng (University of Science and Technology of China) and Shangfei Wang (University of Science and Technology of China)

Non-linear Temporal Subspace Representations for Activity Recognition 2197
Anoop Cherian (MERL, Cambridge MA), Anoop Cherian (ACRV, ANU Canberra), Suvrit Sra (MIT, Cambridge MA), Stephen Gould (ACRV, ANU Canberra), and Richard Hartley (ACRV, ANU Canberra)

Towards Pose Invariant Face Recognition in the Wild 2207
Jian Zhao (National University of Singapore; National University of Defense Technology), Yu Cheng (Nanyang Technological University), Yan Xu (Panasonic R&D Center Singapore), Lin Xiong (Panasonic R&D Center Singapore), Jianshu Li (National University of Singapore), Fang Zhao (National University of Singapore), Karlekar Jayashree (Panasonic R&D Center Singapore), Sugiri Pranata (Panasonic R&D Center Singapore), Shengmei Shen (Panasonic R&D Center Singapore), Junliang Xing (National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences), Shuicheng Yan (National University of Singapore; Qihoo 360 AI Institute), and Jiashi Feng (National University of Singapore)

Unifying Identification and Context Learning for Person Recognition 2217
Qingqiu Huang (CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong), Yu Xiong (CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong), and Dahua Lin (CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong)

Jointly Optimize Data Augmentation and Network Training: Adversarial Data Augmentation in Human Pose Estimation 2226
Xi Peng (Rutgers University), Zhiqiang Tang (Rutgers University), Fei Yang (Facebook), Rogerio S. Feris (IBM T.J. Watson Research Center), and Dimitris Metaxas (Rutgers University)

Wing Loss for Robust Facial Landmark Localisation with Convolutional Neural Networks 2235
Zhen-Hua Feng (University of Surrey, Guildford, UK), Josef Kittler (University of Surrey, Guildford, UK), Muhammad Awais (University of Surrey, Guildford, UK), Patrik Huber (University of Surrey, Guildford, UK), and Xiao-Jun Wu (Jiangnan University, China)

Multiple Granularity Group Interaction Prediction 2246
Taiping Yao (Shanghai Jiao Tong University, Shanghai, China), Minsi Wang (Shanghai Jiao Tong University, Shanghai, China), Bingbing Ni (Shanghai Jiao Tong University, Shanghai, China), Huawei Wei (Shanghai Jiao Tong University, Shanghai, China), and Xiaokang Yang (Shanghai Jiao Tong University, Shanghai, China)

Social GAN: Socially Acceptable Trajectories with Generative Adversarial Networks 2255
Agrim Gupta (Stanford University), Justin Johnson (Stanford University), Li Fei-Fei (Stanford University), Silvio Savarese (Stanford University), and Alexandre Alahi (Stanford University; École Polytechnique Fédérate de Lausanne)

Deep Group-Shuffling Random Walk for Person Re-identification 2265
Yantao Shen (CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong), Hongsheng Li (CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong), Tong Xiao (CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong), Shuai Yi (SenseTime Research), Dapeng Chen (CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong), and Xiaogang Wang (CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong)

Transferable Joint Attribute-Identity Deep Learning for Unsupervised Person Re-identification 2275
Jingya Wang (Queen Mary University of London), Xiatian Zhu (Vision Semantics Ltd.), Shaogang Gong (Queen Mary University of London), and Wei Li (Queen Mary University of London)

Harmonious Attention Network for Person Re-identification 2285
Wei Li (Queen Mary University of London), Xiatian Zhu (Vision Semantics Ltd.), and Shaogang Gong (Queen Mary University of London)

Real-Time Rotation-Invariant Face Detection with Progressive Calibration Networks 2295
Xuepeng Shi (Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology, CAS, Beijing, China), Xuepeng Shi (University of Chinese Academy of Sciences, Beijing, China), Shiguang Shan (Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology, CAS, Beijing, China), Meina Kan (Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology, CAS, Beijing, China), Shuzhe Wu (Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology, CAS, Beijing, China), and Xilin Chen (Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology, CAS, Beijing, China)

Deep Regression Forests for Age Estimation 2304
Wei Shen (Shanghai University, China; Johns Hopkins University), Yilu Guo (Shanghai University, China), Yan Wang (Johns Hopkins University), Kai Zhao (Nankai University, China), Bo Wang (Hikvision Research), and Alan Yuille (Johns Hopkins University)

Weakly-Supervised Deep Convolutional Neural Network Learning for Facial Action Unit Intensity Estimation 2314
Yong Zhang (National Laboratory of Pattern Recognition, CASIA), Yong Zhang (University of Chinese Academy of Sciences), Weiming Dong (National Laboratory of Pattern Recognition, CASIA), Bao-Gang Hu (National Laboratory of Pattern Recognition, CASIA), and Qiang Ji (Rensselaer Polytechnic Institute)

Memory Based Online Learning of Deep Representations from Video Streams 2324
Federico Pernici (MICC - Media Integration and Communcation Center University Of Florence - Italy), Federico Bartoli (MICC - Media Integration and Communcation Center University Of Florence - Italy), Matteo Bruni (MICC - Media Integration and Communcation Center University Of Florence - Italy), and Alberto Del Bimbo (MICC - Media Integration and Communcation Center University Of Florence - Italy)

Efficient and Deep Person Re-identification Using Multi-level Similarity 2335
Yiluan Guo (Singapore University of Technology and Design) and Ngai-Man Cheung (Singapore University of Technology and Design)

Multi-level Fusion Based 3D Object Detection from Monocular Images 2345
Bin Xu (Wuhan University, China) and Zhenzhong Chen (Wuhan University, China)

A Perceptual Measure for Deep Single Image Camera Calibration 2354
Yannick Hold-Geoffroy (Université Laval), Kalyan Sunkavalli (Adobe), Jonathan Eisenmann (Adobe), Matt Fisher (Adobe), Emiliano Gambaretto (Adobe), Sunil Hadap (Adobe), and Jean-Francois Lalonde (Université Laval)

Learning to Generate Time-Lapse Videos Using Multi-stage Dynamic Generative Adversarial Networks 2364
Wei Xiong (Tencent AI Lab; University of Rochester), Wenhan Luo (Tencent AI Lab), Lin Ma (Tencent AI Lab), Wei Liu (Tencent AI Lab), and Jiebo Luo (University of Rochester; Tencent AI Lab)

Document Enhancement Using Visibility Detection 2374
Netanel Kligler (Technion Haifa, Israel), Sagi Katz (Technion Haifa, Israel), and Ayellet Tal (Technion Haifa, Israel)

A Weighted Sparse Sampling and Smoothing Frame Transition Approach for Semantic Fast-Forward First-Person Videos 2383
Michel Silva (Universidade Federal de Minas Gerais (UFMG)), Washington Ramos (Universidade Federal de Minas Gerais (UFMG)), Joao Ferreira (Universidade Federal de Minas Gerais (UFMG)), Felipe Chamone (Universidade Federal de Minas Gerais (UFMG)), Mario Campos (Universidade Federal de Minas Gerais (UFMG)), and Erickson R. Nascimento (Universidade Federal de Minas Gerais (UFMG))

Oral and Spotlight 1-2A

O1-2A: Machine Learning for Computer Vision I

Context Contrasted Feature and Gated Multi-scale Aggregation for Scene Segmentation 2393
Henghui Ding (Nanyang Technological University, Singapore), Xudong Jiang (Nanyang Technological University, Singapore), Bing Shuai (Nanyang Technological University, Singapore), Ai Qun Liu (Nanyang Technological University, Singapore), and Gang Wang (Alibaba Group, Hangzhou, China)

Deep Layer Aggregation 2403
Fisher Yu (UC Berkeley), Dequan Wang (UC Berkeley), Evan Shelhamer (UC Berkeley), and Trevor Darrell (UC Berkeley)

Convolutional Neural Networks with Alternately Updated Clique 2413
Yibo Yang (Peking University), Zhisheng Zhong (Peking University), Tiancheng Shen (Peking University; Peking University), and Zhouchen Lin (Peking University; Shanghai Jiao Tong University)

Practical Block-Wise Neural Network Architecture Generation 2423
Zhao Zhong (Chinese Academy of Sciences; University of Chinese Academy of Sciences), Junjie Yan (SenseTime Research Institute), Wei Wu (SenseTime Research Institute), Jing Shao (SenseTime Research Institute), and Cheng-Lin Liu (University of Chinese Academy of Sciences; CAS Center for Excellence of Brain Science and Intelligence Technology)

S1-2A

xUnit: Learning a Spatial Activation Function for Efficient Image Restoration 2433
Idan Kligvasser (Technion-Israel Institute of Technology, Haifa, Israel), Tamar Rott Shaham (Technion-Israel Institute of Technology, Haifa, Israel), and Tomer Michaeli (Technion-Israel Institute of Technology, Haifa, Israel)

Crafting a Toolchain for Image Restoration by Deep Reinforcement Learning 2443
Ke Yu (The Chinese University of Hong Kong), Chao Dong (SenseTime Research), Liang Lin (SenseTime Research; Sun Yat-sen University), and Chen Change Loy (CUHK - SenseTime Joint Lab, The Chinese University of Hong Kong)

Deformation Aware Image Compression 2453
Tamar Rott Shaham (Technion-Israel Institute of Technology, Haifa, Israel) and Tomer Michaeli (Technion-Israel Institute of Technology, Haifa, Israel)

Distributable Consistent Multi-object Matching 2463
Nan Hu (Stanford University), Qixing Huang (UT Austin), Boris Thibert (UG Alpes), and Leonidas Guibas (Stanford University)

Residual Dense Network for Image Super-Resolution 2472
Yulun Zhang (Northeastern University, Boston, USA), Yapeng Tian (University of Rochester, Rochester, USA), Yu Kong (Northeastern University, Boston, USA), Bineng Zhong (Northeastern University, Boston, USA), and Yun Fu (Northeastern University, Boston, USA; Northeastern University, Boston, USA)

Attentive Generative Adversarial Network for Raindrop Removal from A Single Image 2482
Rui Qian (Peking University, Beijing, P.R. China), Robby T. Tan (National University of Singapore; Yale-NUS College), Wenhan Yang (Peking University, Beijing, P.R. China), Jiajun Su (Peking University, Beijing, P.R. China), and Jiaying Liu (Peking University, Beijing, P.R. China)

FSRNet: End-to-End Learning Face Super-Resolution with Facial Priors 2492
Yu Chen (Nanjing University of Science and Technology; Motovis Inc.), Ying Tai (Youtu Lab, Tencent), Xiaoming Liu (Michigan State University), Chunhua Shen (University of Adelaide), and Jian Yang (Nanjing University of Science and Technology)

Burst Denoising with Kernel Prediction Networks 2502
Ben Mildenhall (UC Berkeley; Google Research), Jonathan T. Barron (Google Research), Jiawen Chen (Google Research), Dillon Sharlet (Google Research), Ren Ng (UC Berkeley), and Robert Carroll (Google Research)

Unsupervised Sparse Dirichlet-Net for Hyperspectral Image Super-Resolution 2511
Ying Qu (The University of Tennessee, Knoxville, TN), Hairong Qi (The University of Tennessee, Knoxville, TN), and Chiman Kwan (Applied Research LLC, Rockville, MD)

Dynamic Scene Deblurring Using Spatially Variant Recurrent Neural Networks 2521
Jiawei Zhang (City University of Hong Kong; SenseTime Research), Jinshan Pan (School of Computer Science and Engineering, Nanjing University of Science and Technology), Jimmy Ren (SenseTime Research), Yibing Song (Tencent AI Lab), Linchao Bao (Tencent AI Lab), Rynson W.H. Lau (City University of Hong Kong), and Ming-Hsuan Yang (University of California, Merced)

Oral and Spotlight 1-2B

O1-2B: 3D Vision II

SPLATNet: Sparse Lattice Networks for Point Cloud Processing 2530
Hang Su (UMass Amherst), Varun Jampani (NVIDIA), Deqing Sun (NVIDIA), Subhransu Maji (UMass Amherst), Evangelos Kalogerakis (UMass Amherst), Ming-Hsuan Yang (UC Merced), and Jan Kautz (NVIDIA)

Surface Networks 2540
Ilya Kostrikov (Courant Institute of Mathematical Sciences, New York University), Zhongshi Jiang (Courant Institute of Mathematical Sciences, New York University), Daniele Panozzo (Courant Institute of Mathematical Sciences, New York University), Denis Zorin (Courant Institute of Mathematical Sciences, New York University), and Joan Bruna (Courant Institute of Mathematical Sciences, New York University; Center for Data Science, New York University)

Self-Supervised Multi-level Face Model Learning for Monocular Reconstruction at Over 250 Hz 2549
Ayush Tewari (MPI Informatics; Saarland Informatics Campus), Michael Zollhöfer (MPI Informatics; Saarland Informatics Campus; Stanford University), Pablo Garrido (MPI Informatics; Saarland Informatics Campus), Florian Bernard (MPI Informatics; Saarland Informatics Campus), Hyeongwoo Kim (MPI Informatics; Saarland Informatics Campus), Patrick Perez (Technicolor), and Christian Theobalt (MPI Informatics; Saarland Informatics Campus)

CodeSLAM - Learning a Compact, Optimisable Representation for Dense Visual SLAM 2560
Michael Bloesch (Imperial College London), Jan Czarnowski (Imperial College London), Ronald Clark (Imperial College London), Stefan Leutenegger (Imperial College London), and Andrew J. Davison (Imperial College London)

S1-2B

SGPN: Similarity Group Proposal Network for 3D Point Cloud Instance Segmentation 2569
Weiyue Wang (University of Southern California Los Angeles, California), Ronald Yu (University of California, San Diego San Diego, California), Qiangui Huang (University of Southern California Los Angeles, California), and Ulrich Neumann (University of Southern California Los Angeles, California)

PlaneNet: Piece-Wise Planar Reconstruction from a Single RGB Image 2579
Chen Liu (Washington University in St. Louis), Jimei Yang (Adobe Research), Duygu Ceylan (Adobe Research), Ersin Yumer (Argo AI), and Yasutaka Furukawa (Simon Fraser University)

Deep Parametric Continuous Convolutional Neural Networks 2589
Shenlong Wang (University of Toronto; Uber Advanced Technologies Group), Simon Suo (University of Waterloo; Advanced Technologies Group), Wei-Chiu Ma (Uber Advanced Technologies Group), Andrei Pokrovsky (Uber Advanced Technologies Group), and Raquel Urtasun (University of Toronto; Uber Advanced Technologies Group)

FeaStNet: Feature-Steered Graph Convolutions for 3D Shape Analysis 2598
Nitika Verma (University Grenoble Alpes, Inria, France), Edmond Boyer (University Grenoble Alpes, Inria, France), and Jakob Verbeek (University Grenoble Alpes, Inria, France)

Image Collection Pop-up: 3D Reconstruction and Clustering of Rigid and Non-rigid Categories 2607
Antonio Agudo (Institut de Robòtica i Informàtica Industrial, Barcelona, Spain), Melcior Pijoan (Institut de Robòtica i Informàtica Industrial, Barcelona, Spain), and Francesc Moreno-Noguer (Institut de Robòtica i Informàtica Industrial, Barcelona, Spain)

Geometry-Aware Learning of Maps for Camera Localization 2616
Samarth Brahmbhatt (Georgia Institute of Technology), Jinwei Gu (NVIDIA), Kihwan Kim (NVIDIA), James Hays (Georgia Institute of Technology), and Jan Kautz (NVIDIA)

Recurrent Slice Networks for 3D Segmentation of Point Clouds 2626
Qiangui Huang (University of Southern California, Los Angeles, California), Weiyue Wang (University of Southern California, Los Angeles, California), and Ulrich Neumann (University of Southern California, Los Angeles, California)

Depth-Based 3D Hand Pose Estimation: From Current Achievements to Future Goals 2636
Shanxin Yuan (Imperial College London), Guillermo Garcia-Hernando (Imperial College London), Björn Stenger (Rakuten Institute of Technology), Gyeongsik Moon (Seoul National University), Ju Yong Chang (Kwangwoon University), Kyoung Mu Lee (Seoul National University), Pavlo Molchanov (NVIDIA), Jan Kautz (NVIDIA), Sina Honari (University of Montreal), Liuhao Ge (Nanyang Technological University), Junsong Yuan (State University of New York at Buffalo), Xinghao Chen (Tsinghua University, China), Guijin Wang (Tsinghua University, China), Fan Yang (Nara Institute of Science and Technology), Kai Akiyama (Nara Institute of Science and Technology), Yang Wu (Nara Institute of Science and Technology), Qingfu Wan (Fudan University), Meysam Madadi (Computer Vision Center), Sergio Escalera (Computer Vision Center; University of Barcelona), Shile Li (Technical University of Munich), Dongheui Lee (Technical University of Munich; German Aerospace Center), Iason Oikonomidis (University of Crete and FORTH), Antonis Argyros (University of Crete and FORTH), and Tae-Kyun Kim (Imperial College London)

SobolevFusion: 3D Reconstruction of Scenes Undergoing Free Non-rigid Motion 2646
Miroslava Slavcheva (Technische Universitat München; Siemens Corporate Technology), Maximilian Baust (Technische Universitat München), and Slobodan Ilic (Technische Universitat München; Siemens Corporate Technology)

AdaDepth: Unsupervised Content Congruent Adaptation for Depth Estimation 2656
Jogendra Nath Kundu (Indian Institute of Science, Bangalore, India), Phani Krishna Uppala (Indian Institute of Science, Bangalore, India), Anuj Pahuja (Indian Institute of Science, Bangalore, India), and R. Venkatesh Babu (Indian Institute of Science, Bangalore, India)

Oral and Spotlight 1-2C

O1-2C: Machine Learning for Computer Vision II

Learning to Find Good Correspondences 2666
Kwang Moo Yi (University of Victoria), Eduard Trulls (Computer Vision Laboratory, École Polytechnique Fédérale de Lausanne), Yuki Ono (Sony Imaging Products & Solutions Inc.), Vincent Lepetit (Institute for Computer Graphics and Vision, Graz University of Technology), Mathieu Salzmann (Computer Vision Laboratory, École Polytechnique Fédérale de Lausanne), and Pascal Fua (Computer Vision Laboratory, École Polytechnique Fédérale de Lausanne)

OATM: Occlusion Aware Template Matching by Consensus Set Maximization 2675
Simon Korman (Weizmann Institute of Science; UCLA Vision Lab), Stefano Soatto (Weizmann Institute of Science; UCLA Vision Lab), and Mark Milam (Northrop Grumman)

Deep Learning of Graph Matching 2684
Andrei Zanfir (Lund University) and Cristian Sminchisescu (Institute of Mathematics of the Romanian Academy)

Unsupervised Discovery of Object Landmarks as Structural Representations 2694
Yuting Zhang (University of Michigan, Ann Arbor), Yijie Guo (University of Michigan, Ann Arbor), Yixin Jin (University of Michigan, Ann Arbor), Yijun Luo (University of Michigan, Ann Arbor), Zhiyuan He (University of Michigan, Ann Arbor), and Honglak Lee (Google Brain; University of Michigan, Ann Arbor)

S1-2C

Quantization and Training of Neural Networks for Efficient Integer-Arithmetic-Only Inference 2704
Benoit Jacob (Google Inc.), Skirmantas Kligys (Google Inc.), Bo Chen (Google Inc.), Menglong Zhu (Google Inc.), Matthew Tang (Google Inc.), Andrew Howard (Google Inc.), Hartwig Adam (Google Inc.), and Dmitry Kalenichenko (Google Inc.)

Lean Multiclass Crowdsourcing 2714
Grant Van Horn (Caltech), Steve Branson (Caltech), Scott Loarie (iNaturalist), Serge Belongie (Cornell Tech), and Pietro Perona (Caltech)

Partial Transfer Learning with Selective Adversarial Networks 2724
Zhangjie Cao (KLiss, MOE; Tsinghua University, China), Mingsheng Long (KLiss, MOE; Tsinghua University, China), Jianmin Wang (KLiss, MOE; Tsinghua University, China), and Michael I. Jordan (University of California, Berkeley, USA)

Self-Supervised Feature Learning by Learning to Spot Artifacts 2733
Simon Jenni (University of Bern, Switzerland) and Paolo Favaro (University of Bern, Switzerland)

LDMNet: Low Dimensional Manifold Regularized Neural Networks 2743
Wei Zhu (Duke University), Qiang Qiu (Duke University), Jiaji Huang (Baidu Silicon Valley AI Lab), Robert Calderbank (Duke University), Guillermo Sapiro (Duke University), and Ingrid Daubechies (Duke University)

CondenseNet: An Efficient DenseNet Using Learned Group Convolutions 2752
Gao Huang (Cornell University), Shichen Liu (Tsinghua University, China), Laurens van der Maaten (Facebook AI Research), and Kilian Q. Weinberger (Cornell University)

Learning Deep Descriptors with Scale-Aware Triplet Networks 2762
Michel Keller (ETH Zurich, Switzerland), Zetao Chen (ETH Zurich, Switzerland), Fabiola Maffra (ETH Zurich, Switzerland), Patrik Schmuck (ETH Zurich, Switzerland), and Margarita Chli (ETH Zurich, Switzerland)

Decoupled Networks 2771
Weiyang Liu (Georgia Institute of Technology), Zhen Liu (Georgia Institute of Technology), Zhiding Yu (NVIDIA), Bo Dai (Georgia Institute of Technology), Rongmei Lin (Emory University), Yisen Wang (Georgia Institute of Technology; Tsinghua University, China), James M. Rehg (Georgia Institute of Technology), and Le Song (Georgia Institute of Technology; Ant Financial)

Deep Adversarial Metric Learning 2780
Yueqi Duan (Tsinghua University, China; State Key Lab of Intelligent Technologies and Systems; Beijing National Research Center for Information Science and Technology), Wenzhao Zheng (Tsinghua University, China), Xudong Lin (Tsinghua University, China), Jiwen Lu (Tsinghua University, China; State Key Lab of Intelligent Technologies and Systems; Beijing National Research Center for Information Science and Technology), and Jie Zhou (Tsinghua University, China; State Key Lab of Intelligent Technologies and Systems; Beijing National Research Center for Information Science and Technology)

Poster Session P1-3

PU-Net: Point Cloud Upsampling Network 2790
Lequan Yu (The Chinese University of Hong Kong; Chinese Academy of Sciences, China), Xianzhi Li (The Chinese University of Hong Kong), Chi-Wing Fu (The Chinese University of Hong Kong; Chinese Academy of Sciences, China), Daniel Cohen-Or (Tel Aviv University), and Pheng-Ann Heng (The Chinese University of Hong Kong; Chinese Academy of Sciences, China)

Real-Time Monocular Depth Estimation Using Synthetic Data with Domain Adaptation via Image Style Transfer 2800
Amir Atapour-Abarghouei (Durham University), Toby P. Breckon (Durham University), and Toby P. Breckon (Durham University)

Learning for Disparity Estimation Through Feature Constancy 2811
Zhengfa Liang (National University of Defense Technology, Changsha, China), Yiliu Feng (National University of Defense Technology, Changsha, China), Yulan Guo (National University of Defense Technology, Changsha, China), Hengzhu Liu (National University of Defense Technology, Changsha, China), Wei Chen (National University of Defense Technology, Changsha, China), Linbo Qiao (National University of Defense Technology, Changsha, China), Li Zhou (National University of Defense Technology, Changsha, China), and Jianfeng Zhang (National University of Defense Technology, Changsha, China)

DeepMVS: Learning Multi-view Stereopsis 2821
Po-Han Huang (University of Illinois, Urbana-Champaign), Kevin Matzen (Facebook), Johannes Kopf (Facebook), Narendra Ahuja (University of Illinois, Urbana-Champaign), and Jia-Bin Huang (Virginia Tech)

Self-Calibrating Polarising Radiometric Calibration 2831
Daniel Teo Guangwei (Information Systems Technology and Design, Singapore University of Technology and Design), Boxin Shi (National Engineering Laboratory for Video Technology, School of EECS, Peking University), Yinqiang Zheng (Digital Content and Media Sciences Research Division, National Institute of Informatics), and Sai-Kit Yeung (Information Systems Technology and Design, Singapore University of Technology and Design)

Coding Kendall's Shape Trajectories for 3D Action Recognition 2840
Amor Ben Tanfous (CRIStAL - Centre de Recherche en Informatique Signal et Automatique de Lille, France), Hassen Drira (CRIStAL - Centre de Recherche en Informatique Signal et Automatique de Lille, France), and Boulbaba Ben Amor (CRIStAL - Centre de Recherche en Informatique Signal et Automatique de Lille, France)

Efficient, Sparse Representation of Manifold Distance Matrices for Classical Scaling 2850
Javier S. Turek (Intel Labs, Hillsboro, Oregon) and Alexander G. Huth (The University of Texas at Austin)

Motion Segmentation by Exploiting Complementary Geometric Models 2859
Xun Xu (National University of Singapore), Loong Fah Cheong (National University of Singapore), and Zhuwen Li (Intel Labs)

Estimation of Camera Locations in Highly Corrupted Scenarios: All About that Base, No Shape Trouble 2868
Gilad Lerman (University of Minnesota) and Yunpeng Shi (University of Minnesota)

4D Human Body Correspondences from Panoramic Depth Maps 2877
Zhong Li (University of Delaware, Newark, DE, USA), Minye Wu (ShanghaiTech University, Shanghai, China), Wangyiteng Zhou (ShanghaiTech University, Shanghai, China), and Jingyi Yu (University of Delaware, Newark, DE, USA; ShanghaiTech University, Shanghai, China)

Reconstructing Thin Structures of Manifold Surfaces by Integrating Spatial Curves 2887
Shiwei Li (The Hong Kong University of Science and Technology), Yao Yao (The Hong Kong University of Science and Technology), Tian Fang (Altizure.com), and Long Quan (The Hong Kong University of Science and Technology)

Multi-view Consistency as Supervisory Signal for Learning Shape and Pose Prediction 2897
Shubham Tulsiani (University of California, Berkeley), Alexei A. Efros (University of California, Berkeley), and Jitendra Malik (University of California, Berkeley)

Probabilistic Plant Modeling via Multi-view Image-to-Image Translation 2906
Takahiro Isokane (Osaka University), Fumio Okura (Osaka University; JST PRESTO), Ayaka Ide (Osaka University), Yasuyuki Matsushita (Osaka University), and Yasushi Yagi (Osaka University)

Deep Marching Cubes: Learning Explicit Surface Representations 2916
Yiyi Liao (Autonomous Vision Group, MPI for Intelligent Systems Tübingen; Institute of Cyber-Systems and Control, Zhejiang University), Simon Donné (Autonomous Vision Group, MPI for Intelligent Systems Tübingen; imec - IPI - Ghent University), and Andreas Geiger (Autonomous Vision Group, MPI for Intelligent Systems Tübingen; CVG Group, ETH Zurich)

Tags2Parts: Discovering Semantic Regions from Shape Tags 2926
Sanjeev Muralikrishnan (IIT Bombay), Vladimir G. Kim (Adobe Research), and Siddhartha Chaudhuri (IIT Bombay; Adobe Research)

Uncalibrated Photometric Stereo Under Natural Illumination 2936
Zhipeng Mo (Singapore University of Technology and Design), Boxin Shi (Peking University), Feng Lu (Beihang University), Sai-Kit Yeung (Singapore University of Technology and Design), and Yasuyuki Matsushita (Osaka University)

Robust Depth Estimation from Auto Bracketed Images 2946
Sunghoon Im (Korea Advanced Institute of Science and Technology (KAIST), Republic of Korea), Hae-Gon Jeon (Korea Advanced Institute of Science and Technology (KAIST), Republic of Korea), and In So Kweon (Korea Advanced Institute of Science and Technology (KAIST), Republic of Korea)

Free Supervision from Video Games 2955
Philipp Krahenbuhl (UT Austin)

Planar Shape Detection at Structural Scales 2965
Hao Fang (Inria – Universite Côte d'Azur), Florent Lafarge (Inria – Universite Côte d'Azur), and Mathieu Desbrun (Caltech)

Pix3D: Dataset and Methods for Single-Image 3D Shape Modeling 2974
Xingyuan Sun (Massachusetts Institute of Technology; Shanghai Jiao Tong University), Jiajun Wu (Massachusetts Institute of Technology), Xiuming Zhang (Massachusetts Institute of Technology), Zhoutong Zhang (Massachusetts Institute of Technology), Chengkai Zhang (Massachusetts Institute of Technology), Tianfan Xue (Google Research), Joshua B. Tenenbaum (Massachusetts Institute of Technology), and William T. Freeman (Massachusetts Institute of Technology; Google Research)

Camera Pose Estimation with Unknown Principal Point 2984
Viktor Larsson (Lund University Lund, Sweden), Zuzana Kukelova (FEE, Czech Technical University Prague, Czech Republic), and Yinqiang Zheng (National Institute of Informatics Tokyo, Japan)

Inverse Composition Discriminative Optimization for Point Cloud Registration 2993
Jayakorn Vongkulbhisal (ISR - IST, Universidade de Lisboa, Lisboa, Portugal; Carnegie Mellon University, Pittsburgh, PA, USA), Beñat Irastorza Ugalde (Carnegie Mellon University, Pittsburgh, PA, USA), Fernando De la Torre (Carnegie Mellon University, Pittsburgh, PA, USA; Facebook Inc., Menlo Park, CA, USA), and João P. Costeira (ISR - IST, Universidade de Lisboa, Lisboa, Portugal)

SurfConv: Bridging 3D and 2D Convolution for RGBD Images 3002
Hang Chu (University of Toronto; Vector Institute), Wei-Chiu Ma (Uber ATG), Kaustav Kundu (University of Toronto; Vector Institute), Raquel Urtasun (University of Toronto; Vector Institute; Uber ATG), and Sanja Fidler (University of Toronto; Vector Institute)

A Fast Resection-Intersection Method for the Known Rotation Problem 3012
Qianggong Zhang (The University of Adelaide, Adelaide, SA, Australia), Tat-Jun Chin (The University of Adelaide, Adelaide, SA, Australia), and Huu Minh Le (The University of Adelaide, Adelaide, SA, Australia)

3D Pose Estimation and 3D Model Retrieval for Objects in the Wild 3022
Alexander Grabner (Institute of Computer Graphics and Vision, Graz University of Technology, Austria), Peter M. Roth (Institute of Computer Graphics and Vision, Graz University of Technology, Austria), and Vincent Lepetit (Institute of Computer Graphics and Vision, Graz University of Technology, Austria; Laboratoire Bordelais de Recherche en Informatique, University of Bordeaux, France)

Structure from Recurrent Motion: From Rigidity to Recurrency 3032
Xiu Li (Tsinghua University, China; Carnegie Mellon University), Hongdong Li (Carnegie Mellon University; Australian National University), Hanbyul Joo (Carnegie Mellon University), Yebin Liu (Tsinghua University, China), and Yaser Sheikh (Carnegie Mellon University)

Learning Patch Reconstructability for Accelerating Multi-view Stereo 3041
Alex Poms (Carnegie Mellon University, Pittsburgh, PA), Chenglei Wu (Facebook Reality Labs, Pittsburgh, Pittsburgh, PA), Shoou-I Yu (Facebook Reality Labs, Pittsburgh, Pittsburgh, PA), and Yaser Sheikh (Facebook Reality Labs, Pittsburgh, Pittsburgh, PA)

Progressively Complementarity-Aware Fusion Network for RGB-D Salient Object Detection 3051
Hao Chen (City University of Hong Kong, Kowloon, Hong Kong) and Youfu Li (City University of Hong Kong, Kowloon, Hong Kong)

Pixels, Voxels, and Views: A Study of Shape Representations for Single View 3D Object Shape Prediction 3061
Daeyun Shin (University of California, Irvine), Charless C. Fowlkes (University of California, Irvine), and Derek Hoiem (University of Illinois, Urbana-Champaign)

Learning Dual Convolutional Neural Networks for Low-Level Vision 3070
Jinshan Pan (Nanjing University of Science and Technology), Sifei Liu (NVIDIA), Deqing Sun (NVIDIA), Jiawei Zhang (City University of Hong Kong), Yang Liu (Dalian University of Technology), Jimmy Ren (SenseTime Research), Zechao Li (Nanjing University of Science and Technology), Jinhui Tang (Nanjing University of Science and Technology), Huchuan Lu (Dalian University of Technology), Yu-Wing Tai (Tencent Youtu Lab), and Ming-Hsuan Yang (UC Merced)

Defocus Blur Detection via Multi-stream Bottom-Top-Bottom Fully Convolutional Network 3080
Wenda Zhao (Dalian University of Technology, China), Fan Zhao (Dalian Institute of Chemical Physics, Chinese Academy of Sciences, China), Dong Wang (Dalian University of Technology, China), and Huchuan Lu (Dalian University of Technology, China)

PiCANet: Learning Pixel-Wise Contextual Attention for Saliency Detection 3089
Nian Liu (Northwestern Polytechincal University), Junwei Han (Northwestern Polytechincal University), and Ming-Hsuan Yang (University of California, Merced; Google Cloud)

Curve Reconstruction via the Global Statistics of Natural Curves 3099
Ehud Barnea (Ben-Gurion University Beer-Sheva, Israel) and Ohad Ben-Shahar (Ben-Gurion University Beer-Sheva, Israel)

What do Deep Networks Like to See? 3108
Sebastian Palacio (German Research Center for Artificial Intelligence (DFKI) TU Kaiserslautern), Joachim Folz (German Research Center for Artificial Intelligence (DFKI) TU Kaiserslautern), Jörn Hees (German Research Center for Artificial Intelligence (DFKI) TU Kaiserslautern), Federico Raue (German Research Center for Artificial Intelligence (DFKI) TU Kaiserslautern), Damian Borth (German Research Center for Artificial Intelligence (DFKI) TU Kaiserslautern), and Andreas Dengel (German Research Center for Artificial Intelligence (DFKI) TU Kaiserslautern)

Zero-Shot Super-Resolution Using Deep Internal Learning 3118
Assaf Shocher (The Weizmann Institute of Science, Israel), Nadav Cohen (School of Mathematics, Institute for Advanced Study, Princeton, New Jersey), and Michal Irani (The Weizmann Institute of Science, Israel)

Detect Globally, Refine Locally: A Novel Approach to Saliency Detection 3127
Tiantian Wang (Dalian University of Technology), Lihe Zhang (Dalian University of Technology), Shuo Wang (Dalian University of Technology), Huchuan Lu (Dalian University of Technology), Gang Yang (Northeastern University, China), Xiang Ruan (Tiwaki Co., Ltd), and Ali Borji (University of Central Florida, USA)

Beyond the Pixel-Wise Loss for Topology-Aware Delineation 3136
Agata Mosinska (Computer Vision Laboratory, École Polytechnique Fédérale de Lausanne (EPFL)), Pablo Marquez-Neila (Computer Vision Laboratory, École Polytechnique Fédérale de Lausanne (EPFL); ARTORG Center for Biomedical Engineering Research, University of Bern), Mateusz Kozinski (Computer Vision Laboratory, École Polytechnique Fédérale de Lausanne (EPFL)), and Pascal Fua (Computer Vision Laboratory, École Polytechnique Fédérale de Lausanne (EPFL))

KIPPI: KInetic Polygonal Partitioning of Images 3146
Jean-Philippe Bauchet (Inria - Université Côte d'Azur) and Florent Lafarge (Inria - Université Côte d'Azur)

Image Blind Denoising with Generative Adversarial Network Based Noise Modeling 3155
Jingwen Chen (Sun Yat-sen University, China), Jiawei Chen (Sun Yat-sen University, China), Hongyang Chao (Sun Yat-sen University, China), and Ming Yang (CVTE Research, Guangzhou, P. R. China)

Multi-Scale Weighted Nuclear Norm Image Restoration 3165
Noam Yair (Technion - Israel Institute of Technology) and Tomer Michaeli (Technion - Israel Institute of Technology)

MoNet: Moments Embedding Network 3175
Mengran Gou (Northeastern University, Boston), Fei Xiong (USC), Octavia Camps (Northeastern University, Boston), and Mario Sznaier (Northeastern University, Boston)

Active Fixation Control to Predict Saccade Sequences 3184
Calden Wloka (York University, Toronto, Canada), Iuliia Kotseruba (York University, Toronto, Canada), and John K. Tsotsos (York University, Toronto, Canada)

Densely Connected Pyramid Dehazing Network 3194
He Zhang (Rutgers University, Piscataway) and Vishal M. Patel (Rutgers University, Piscataway)

Universal Denoising Networks : A Novel CNN Architecture for Image Denoising 3204
Stamatios Lefkimmiatis (Skolkovo Institute of Science and Technology (Skoltech))

Learning Convolutional Networks for Content-Weighted Image Compression 3214
Mu Li (The Hong Kong Polytechnic University), Wangmeng Zuo (Harbin Institute of Technology), Shuhang Gu (The Hong Kong Polytechnic University), Debin Zhao (Harbin Institute of Technology), and David Zhang (The Hong Kong Polytechnic University)

Deep Video Super-Resolution Network Using Dynamic Upsampling Filters Without Explicit Motion Compensation 3224
Younghyun Jo (Yonsei University), Seoung Wug Oh (Yonsei University), Jaeyeon Kang (Yonsei University), and Seon Joo Kim (Yonsei University)

Erase or Fill? Deep Joint Recurrent Rain Removal and Reconstruction in Videos 3233
Jiaying Liu (Peking University, Beijing, P.R. China), Wenhan Yang (Peking University, Beijing, P.R. China), Shuai Yang (Peking University, Beijing, P.R. China), and Zongming Guo (Peking University, Beijing, P.R. China)

Flow Guided Recurrent Neural Encoder for Video Salient Object Detection 3243
Guanbin Li (Sun Yat-sen University), Yuan Xie (Sun Yat-sen University), Tianhao Wei (Zhejiang University), Keze Wang (Sun Yat-sen University), and Liang Lin (Sun Yat-sen University; SenseTime Group Limited)

Gated Fusion Network for Single Image Dehazing 3253
Wenqi Ren (State Key Laboratory of Information Security (SKLOIS), IIE, CAS), Lin Ma (Tencent AI Lab), Jiawei Zhang (City University of Hong Kong), Jinshan Pan (Nanjing University of Science and Technology), Xiaochun Cao (State Key Laboratory of Information Security (SKLOIS), IIE, CAS; School of Cyber Security, University of Chinese Academy of Sciences), Wei Liu (Tencent AI Lab), and Ming-Hsuan Yang (University of California, Merced)

Learning a Single Convolutional Super-Resolution Network for Multiple Degradations 3262
Kai Zhang (Harbin Institute of Technology, Harbin, China; The Hong Kong Polytechnic University, Hong Kong, China; DAMO Academy, Alibaba Group), Wangmeng Zuo (Harbin Institute of Technology, Harbin, China), and Lei Zhang (The Hong Kong Polytechnic University, Hong Kong, China)

Non-blind Deblurring: Handling Kernel Uncertainty with CNNs 3272
Subeesh Vasu (Indian Institute of Technology Madras), Venkatesh Reddy Maligireddy (Indian Institute of Technology Madras), and A. N. Rajagopalan (Indian Institute of Technology Madras)

Boundary Flow: A Siamese Network that Predicts Boundary Motion Without Training on Motion 3282
Peng Lei (Oregon State University Corvallis, OR, USA), Fuxin Li (Oregon State University Corvallis, OR, USA), and Sinisa Todorovic (Oregon State University Corvallis, OR, USA)

Learning to See in the Dark 3291
Chen Chen (UIUC), Qifeng Chen (Intel Labs), Jia Xu (Intel Labs), and Vladlen Koltun (Intel Labs)

BPGrad: Towards Global Optimality in Deep Learning via Branch and Pruning 3301
Ziming Zhang (Mitsubishi Electric Research Laboratories), Yuanwei Wu (University of Kansas), and Guanghui Wang (University of Kansas)

Perturbative Neural Networks 3310
Felix Juefei-Xu (Carnegie Mellon University), Vishnu Naresh Boddeti (Michigan State University), and Marios Savvides (Carnegie Mellon University)

Unsupervised Correlation Analysis 3319
Yedid Hoshen (Facebook AI Research Facebook AI Research and Tel Aviv University) and Lior Wolf (Facebook AI Research Facebook AI Research and Tel Aviv University)

A Biresolution Spectral Framework for Product Quantization 3329
Lopamudra Mukherjee (University of Wisconsin-Whitewater), Jiming Peng (University of Houston), Sathya N. Ravi (University of Wisconsin-Madison), and Vikas Singh (University of Wisconsin-Madison)

Domain Adaptive Faster R-CNN for Object Detection in the Wild 3339
Yuhua Chen (Computer Vision Lab, ETH Zurich), Wen Li (ETH Zurich), Christos Sakaridis (Computer Vision Lab, ETH Zurich), Dengxin Dai (Computer Vision Lab, ETH Zurich), and Luc Van Gool (ETH Zurich; KU Leuven)

Low-Shot Learning with Large-Scale Diffusion 3349
Matthijs Douze (Facebook AI Research), Arthur Szlam (Facebook AI Research), Bharath Hariharan (Facebook AI Research; Cornell University), and Hervé Jégou (Facebook AI Research)

Joint Pose and Expression Modeling for Facial Expression Recognition 3359
Feifei Zhang (Jiangsu University, China; Institute of Automation, CAS), Tianzhu Zhang (Institute of Automation, CAS; University of Chinese Academy of Sciences), Qirong Mao (Jiangsu University, China), and Changsheng Xu (Institute of Automation, CAS; University of Chinese Academy of Sciences)

Lightweight Probabilistic Deep Networks 3369
Jochen Gast (TU Darmstadt) and Stefan Roth (TU Darmstadt)

Adversarially Learned One-Class Classifier for Novelty Detection 3379
Mohammad Sabokrou (Institute for Research in Fundamental Sciences), Mohammad Khalooei (Amirkabir University of Technology), Mahmood Fathy (Institute for Research in Fundamental Sciences), and Ehsan Adeli (Stanford University)

Defense Against Universal Adversarial Perturbations 3389
Naveed Akhtar (The University of Western Australia), Jian Liu (The University of Western Australia), and Ajmal Mian (The University of Western Australia)

Disentangling Factors of Variation by Mixing Them 3399
Matthias Zwicker (University of Maryland, USA), Qiyang Hu (University of Bern, Switzerland), Attila Szabó (University of Bern, Switzerland), Tiziano Portenier (University of Bern, Switzerland), and Paolo Favaro (University of Bern, Switzerland)

Deformable GANs for Pose-Based Human Image Generation 3408
Aliaksandr Siarohin (DISI, University of Trento, Italy), Enver Sangineto (DISI, University of Trento, Italy), Stéphane Lathuilière (Inria Grenoble Rhone-Alpes, France), and Nicu Sebe (DISI, University of Trento, Italy)

Hierarchical Recurrent Attention Networks for Structured Online Maps 3417
Namdar Homayounfar (Uber ATG Toronto), Wei-Chiu Ma (Uber ATG Toronto), Shrinidhi Kowshika Lakshmikanth (Uber ATG Toronto), and Raquel Urtasun (Uber ATG Toronto)

Sliced Wasserstein Distance for Learning Gaussian Mixture Models 3427
Soheil Kolouri (HRL Laboratories, LLC), Gustavo K. Rohde (University of Virginia), and Heiko Hoffmann (HRL Laboratories, LLC)

Aligning Infinite-Dimensional Covariance Matrices in Reproducing Kernel Hilbert Spaces for Domain Adaptation 3437
Zhen Zhang (Washington University in St. Louis), Mianzhi Wang (Washington University in St. Louis), Yan Huang (Washington University in St. Louis), and Arye Nehorai (Washington University in St. Louis)

CLEAR: Cumulative LEARning for One-Shot One-Class Image Recognition 3446
Jedrzej Kozerawski (University of California, Santa Barbara) and Matthew Turk (University of California, Santa Barbara)

Local and Global Optimization Techniques in Graph-Based Clustering 3456
Daiki Ikami (The University of Tokyo, Japan), Toshihiko Yamasaki (The University of Tokyo, Japan), and Kiyoharu Aizawa (The University of Tokyo, Japan)

Multi-task Learning by Maximizing Statistical Dependence 3465
Youssef A. Mejjati (University of Bath), Darren Cosker (University of Bath), and Kwang In Kim (University of Bath)

Robust Classification with Convolutional Prototype Learning 3474
Hong-Ming Yang (NLPR, Institute of Automation, Chinese Academy of Sciences, Beijing, P.R. China; University of Chinese Academy of Sciences, Beijing, P.R. China), Xu-Yao Zhang (NLPR, Institute of Automation, Chinese Academy of Sciences, Beijing, P.R. China; University of Chinese Academy of Sciences, Beijing, P.R. China), Fei Yin (NLPR, Institute of Automation, Chinese Academy of Sciences, Beijing, P.R. China; University of Chinese Academy of Sciences, Beijing, P.R. China), and Cheng-Lin Liu (NLPR, Institute of Automation, Chinese Academy of Sciences, Beijing, P.R. China; University of Chinese Academy of Sciences, Beijing, P.R. China; CAS Center for Excellence of Brain Science and Intelligence Technology, Beijing, P.R. China)

Generative Modeling Using the Sliced Wasserstein Distance 3483
Ishan Deshpande (University of Illinois Urbana-Champaign), Ziyu Zhang (Snap Inc. Los Angeles), and Alexander Schwing (University of Illinois Urbana-Champaign)

Learning Time/Memory-Efficient Deep Architectures with Budgeted Super Networks 3492
Tom Véniat (Sorbonne Université) and Ludovic Denoyer (Sorbonne Université; Criteo Research)

Cross-View Image Synthesis Using Conditional GANs 3501
Krishna Regmi (Center for Research in Computer Vision, University of Central Florida) and Ali Borji (Center for Research in Computer Vision, University of Central Florida)

Sparse, Smart Contours to Represent and Edit Images 3511
Tali Dekel (Google Research), Chuang Gan (MIT-Watson AI Lab), Dilip Krishnan (Google Research), Ce Liu (Google Research), and William T. Freeman (Google Research; MIT-CSAIL)

Anticipating Traffic Accidents with Adaptive Loss and Large-Scale Incident DB 3521
Tomoyuki Suzuki (National Institute of Advanced Industrial Science and Technology (AIST); Keio University), Hirokatsu Kataoka (National Institute of Advanced Industrial Science and Technology (AIST)), Yoshimitsu Aoki (Keio University), and Yutaka Satoh (National Institute of Advanced Industrial Science and Technology (AIST))

A Minimalist Approach to Type-Agnostic Detection of Quadrics in Point Clouds 3530
Tolga Birdal (Technische Universitat München, Germany; Siemens AG, Munich, Germany), Benjamin Busam (Technische Universitat München, Germany; Framos AG, Munich, Germany), Nassir Navab (Technische Universitat München, Germany; Johns Hopkins University, USA), Slobodan Ilic (Technische Universitat München, Germany; Siemens AG, Munich, Germany), and Peter Sturm (INRIA, France)

Facelet-Bank for Fast Portrait Manipulation 3541
Ying-Cong Chen (The Chinese University of Hong Kong), Huaijia Lin (The Chinese University of Hong Kong), Michelle Shu (Johns Hopkins University), Ruiyu Li (The Chinese University of Hong Kong), Xin Tao (The Chinese University of Hong Kong), Xiaoyong Shen (Tencent Youtu Lab), Yangang Ye (Tencent Youtu Lab), and Jiaya Jia (The Chinese University of Hong Kong; Tencent Youtu Lab)

Visual to Sound: Generating Natural Sound for Videos in the Wild 3550
Yipin Zhou (University of North Carolina at Chapel Hill), Zhaowen Wang (Adobe Research), Chen Fang (Adobe Research), Trung Bui (Adobe Research), and Tamara L. Berg (University of North Carolina at Chapel Hill)

Oral and Spotlight 2-1A

3D-RCNN: Instance-Level 3D Object Reconstruction via Render-and-Compare 3559
Abhijit Kundu (Georgia Institute of Technology), Yin Li (Georgia Institute of Technology; Carnegie Mellon University), and James M. Rehg (Georgia Institute of Technology)

Fast and Furious: Real Time End-to-End 3D Detection, Tracking and Motion Forecasting with a Single Convolutional Net 3569
Wenjie Luo (Uber Advanced Technologies Group University of Toronto), Bin Yang (Uber Advanced Technologies Group University of Toronto), and Raquel Urtasun (Uber Advanced Technologies Group University of Toronto)

An Analysis of Scale Invariance in Object Detection - SNIP 3578
Bharat Singh (University of Maryland, College Park) and Larry S. Davis (University of Maryland, College Park)

O2-1A: Object Recognition & Scene Understanding II

Relation Networks for Object Detection 3588
Han Hu (Microsoft Research Asia), Jiayuan Gu (Peking University), Zheng Zhang (Microsoft Research Asia), Jifeng Dai (Microsoft Research Asia), and Yichen Wei (Microsoft Research Asia)

S2-1A

Zero-Shot Sketch-Image Hashing 3598
Yuming Shen (University of East Anglia, Norwich), Li Liu (Inception Institute of Artificial Intelligence (IIAI), Abu Dhabi, UAE), Fumin Shen (University of Electronic Science and Technology of China, China), and Ling Shao (University of East Anglia, Norwich; Inception Institute of Artificial Intelligence (IIAI), Abu Dhabi, UAE)

VizWiz Grand Challenge: Answering Visual Questions from Blind People 3608
Danna Gurari (University of Texas at Austin), Qing Li (University of Science and Technology of China), Abigale J. Stangl (University of Colorado Boulder), Anhong Guo (Carnegie Mellon University), Chi Lin (University of Texas at Austin), Kristen Grauman (University of Texas at Austin), Jiebo Luo (University of Rochester), and Jeffrey P. Bigham (Carnegie Mellon University)

Divide and Grow: Capturing Huge Diversity in Crowd Images with Incrementally Growing CNN 3618
Deepak Babu Sam (Indian Institute of Science Bangalore, India), Neeraj N Sajjan (Indian Institute of Science Bangalore, India), R. Venkatesh Babu (Indian Institute of Science Bangalore, India), and Mukundhan Srinivasan (NVIDIA Bangalore, India)

Structured Set Matching Networks for One-Shot Part Labeling 3627
Jonghyun Choi (Allen Institute for Artificial Intelligence), Jayant Krishnamurthy (Semantic Machines), Aniruddha Kembhavi (Allen Institute for Artificial Intelligence), and Ali Farhadi (Allen Institute for Artificial Intelligence; University of Washington)

Self-Supervised Learning of Geometrically Stable Features Through Probabilistic Introspection 3637
David Novotny (University of Oxford; Computer Vision Group NAVER LABS Europe), Samuel Albanie (University of Oxford), Diane Larlus (Computer Vision Group NAVER LABS Europe), and Andrea Vedaldi (University of Oxford)

Link and Code: Fast Indexing with Graphs and Compact Regression Codes 3646
Matthijs Douze (Facebook AI Research), Alexandre Sablayrolles (Facebook AI Research; Inria), and Hervé Jégou (Facebook AI Research)

Textbook Question Answering Under Instructor Guidance with Memory Networks 3655
Juzheng Li (Tsinghua Lab of Brain and Intelligence), Hang Su (Tsinghua Lab of Brain and Intelligence), Jun Zhu (Tsinghua Lab of Brain and Intelligence), Siyu Wang (Tsinghua University, China), and Bo Zhang (Tsinghua Lab of Brain and Intelligence)

Unsupervised Deep Generative Adversarial Hashing Network 3664
Kamran Ghasedi Dizaji (University of Pittsburgh, PA, USA), Feng Zheng (University of Pittsburgh, PA, USA), Najmeh Sadoughi Nourabadi (University of Pittsburgh, PA, USA), Yanhua Yang (Xidian University, Xi'an, Shanxi, China), Cheng Deng (Xidian University, Xi'an, Shanxi, China), and Heng Huang (University of Pittsburgh, PA, USA)

Vision-and-Language Navigation: Interpreting Visually-Grounded Navigation Instructions in Real Environments 3674
Peter Anderson (Australian National University), Qi Wu (University of Adelaide), Damien Teney (University of Adelaide), Jake Bruce (Queensland University of Technology), Mark Johnson (Macquarie University), Niko Sünderhauf (Queensland University of Technology), Ian Reid (University of Adelaide), Stephen Gould (Australian National University), and Anton van den Hengel (University of Adelaide)

DenseASPP for Semantic Segmentation in Street Scenes 3684
Maoke Yang (DeepMotion), Kun Yu (DeepMotion), Chi Zhang (DeepMotion), Zhiwei Li (DeepMotion), and Kuiyuan Yang (DeepMotion)

Oral and Spotlight 2-1B

O2-1B: Machine Learning for Computer Vision III

Efficient Optimization for Rank-Based Loss Functions 3693
Pritish Mohapatra (IIIT Hyderabad), Michal RolÃ-nek (MPI Tubingen), C.V. Jawahar (IIIT Hyderabad), Vladimir Kolmogorov (IST Austria), and M. Pawan Kumar (University of Oxford, Alan Turing Institute)

Wasserstein Introspective Neural Networks 3702
Kwonjoon Lee (University of California San Diego), Weijian Xu (University of California San Diego), Fan Fan (University of California San Diego), and Zhuowen Tu (University of California San Diego)

Taskonomy: Disentangling Task Transfer Learning 3712
Amir R. Zamir (Stanford University; University of California, Berkeley), Alexander Sax Sax (Stanford University), William Shen Shen (Stanford University), Leonidas Guibas Guibas (Stanford University), Jitendra Malik Malik (University of California, Berkeley), and Silvio Savarese Savarese (Stanford University)

Maximum Classifier Discrepancy for Unsupervised Domain Adaptation 3723
Kuniaki Saito (The University of Tokyo), Kohei Watanabe (The University of Tokyo), Yoshitaka Ushiku (The University of Tokyo), and Tatsuya Harada (The University of Tokyo; RIKEN)

S2-1B

Unsupervised Feature Learning via Non-parametric Instance Discrimination 3733
Zhirong Wu (UC Berkeley / ICSI; Chinese University of Hong Kong), Yuanjun Xiong (Chinese University of Hong Kong; Amazon Rekognition), Stella X. Yu (UC Berkeley / ICSI), and Dahua Lin (Chinese University of Hong Kong)

Multi-task Adversarial Network for Disentangled Feature Learning 3743
Yang Liu (University of Cambridge), Zhaowen Wang (Adobe Research), Hailin Jin (Adobe Research), and Ian Wassell (University of Cambridge)

Learning from Synthetic Data: Addressing Domain Shift for Semantic Segmentation 3752
Swami Sankaranarayanan (UMIACS, University of Maryland, College Park, MD), Yogesh Balaji (UMIACS, University of Maryland, College Park, MD), Arpit Jain (GE Global Research, Niskayuna, NY), Ser Nam Lim (GE Global Research, Niskayuna, NY; Avitas Systems, GE Venture, Boston MA.), and Rama Chellappa (UMIACS, University of Maryland, College Park, MD)

Empirical Study of the Topology and Geometry of Deep Networks 3762
Alhussein Fawzi (University of California, Los Angeles, USA), Seyed-Mohsen Moosavi-Dezfooli (École Polytechnique Fédérale de Lausanne, Switzerland), Pascal Frossard (École Polytechnique Fédérale de Lausanne, Switzerland), and Stefano Soatto (University of California, Los Angeles, USA)

Boosting Domain Adaptation by Discovering Latent Domains 3771
Massimiliano Mancini (sapienza University of Rome; Fondazione Bruno Kessler), Lorenzo Porzi (Mapillary Research), Samuel Rota Bulò (Mapillary Research), Barbara Caputo (sapienza University of Rome; Italian Institute of Technology), and Elisa Ricci (Fondazione Bruno Kessler; University of Trento)

Shape from Shading Through Shape Evolution 3781
Dawei Yang (University of Michigan) and Jia Deng (University of Michigan)

Weakly Supervised Instance Segmentation Using Class Peak Response 3791
Yanzhao Zhou (University of Chinese Academy of Sciences), Yi Zhu (University of Chinese Academy of Sciences), Qixiang Ye (University of Chinese Academy of Sciences), Qiang Qiu (Duke University), and Jianbin Jiao (University of Chinese Academy of Sciences)

Collaborative and Adversarial Network for Unsupervised Domain Adaptation 3801
Weichen Zhang (The University of Sydney), Wanli Ouyang (The University of Sydney), Wen Li (Computer Vision Laboratory, ETH Zurich), and Dong Xu (The University of Sydney)

Environment Upgrade Reinforcement Learning for Non-differentiable Multi-stage Pipelines 3810
Shuqin Xie (Shanghai Jiao Tong University, China), Zitian Chen (Fudan University, China), Chao Xu (Shanghai Jiao Tong University, China), and Cewu Lu (Shanghai Jiao Tong University, China)

Teaching Categories to Human Learners with Visual Explanations 3820
Oisin Mac Aodha (California Institute of Technology), Shihan Su (California Institute of Technology), Yuxin Chen (California Institute of Technology), Pietro Perona (California Institute of Technology), and Yisong Yue (California Institute of Technology)

Oral and Spotlight 2-1C

O2-1C: 3D Vision III

Density Adaptive Point Set Registration 3829
Felix Järemo Lawin (Linköping University, Sweden), Martin Danelljan (Linköping University, Sweden), Fahad Shahbaz Khan (Linköping University, Sweden), Per-Erik Forssén (Linköping University, Sweden), and Michael Felsberg (Linköping University, Sweden)

Left-Right Comparative Recurrent Model for Stereo Matching 3838
Zequn Jie (Tencent AI Lab), Pengfei Wang (National University of Singapore), Yonggen Ling (Tencent AI Lab), Bo Zhao (University of British Columbia), Yunchao Wei (University of Illinois Urbana-Champaign), Jiashi Feng (National University of Singapore), and Wei Liu (Tencent AI Lab)

Im2Pano3D: Extrapolating 360° Structure and Semantics Beyond the Field of View 3847
Shuran Song (Princeton University), Andy Zeng (Princeton University), Angel X. Chang (Princeton University), Manolis Savva (Princeton University), Silvio Savarese (Stanford University), and Thomas Funkhouser (Princeton University)

Polarimetric Dense Monocular SLAM 3857
Luwei Yang (Simon Fraser University), Feitong Tan (Simon Fraser University), Ao Li (Simon Fraser University), Zhaopeng Cui (Simon Fraser University; ETH Zurich), Yasutaka Furukawa (Simon Fraser University), and Ping Tan (Simon Fraser University)

S2-1C

A Unifying Contrast Maximization Framework for Event Cameras, with Applications to Motion, Depth, and Optical Flow Estimation 3867
Guillermo Gallego (University of Zurich and ETH Zurich), Henri Rebecq (University of Zurich and ETH Zurich), and Davide Scaramuzza (University of Zurich and ETH Zurich)

Modeling Facial Geometry Using Compositional VAEs 3877
Timur Bagautdinov (École Polytechnique Fédérale de Lausanne), Chenglei Wu (Facebook Reality Labs, Pittsburgh), Jason Saragih (Facebook Reality Labs, Pittsburgh), Pascal Fua (École Polytechnique Fédérale de Lausanne), and Yaser Sheikh (Facebook Reality Labs, Pittsburgh)

Tangent Convolutions for Dense Prediction in 3D 3887
Maxim Tatarchenko (University of Freiburg), Jaesik Park (Intel Labs), Vladlen Koltun (Intel Labs), and Qian-Yi Zhou (Intel Labs)

RayNet: Learning Volumetric 3D Reconstruction with Ray Potentials 3897
Despoina Paschalidou (Autonomous Vision Group, MPI for Intelligent Systems Tübingen; Max Planck ETH Center for Learning Systems), Ali Osman Ulusoy (Microsoft), Carolin Schmitt (Autonomous Vision Group, MPI for Intelligent Systems Tübingen), Luc van Gool (Computer Vision Lab, ETH Zurich & KU Leuven), and Andreas Geiger (Autonomous Vision Group, MPI for Intelligent Systems Tübingen; Computer Vision and Geometry Group, ETH Zurich)

Neural 3D Mesh Renderer 3907
Hiroharu Kato (The University of Tokyo), Yoshitaka Ushiku (The University of Tokyo), and Tatsuya Harada (The University of Tokyo; RIKEN)

Structured Attention Guided Convolutional Neural Fields for Monocular Depth Estimation 3917
Dan Xu (Multimedia and Human Understanding Group, University of Trento), Wei Wang (Multimedia and Human Understanding Group, University of Trento), Hao Tang (Multimedia and Human Understanding Group, University of Trento), Hong Liu (Key Laboratory of Machine Perception, Shenzhen Graduate School, Peking University), Nicu Sebe (Multimedia and Human Understanding Group, University of Trento), and Elisa Ricci (Multimedia and Human Understanding Group, University of Trento; Technologies of Vision Group, Fondazione Bruno Kessler)

Automatic 3D Indoor Scene Modeling from Single Panorama 3926
Yang Yang (University of Delaware), Shi Jin (ShanghaiTech University), Ruiyang Liu (ShanghaiTech University), Sing Bing Kang (Microsoft Research), and Jingyi Yu (ShanghaiTech University; Plex-VR Inc.)

Extreme 3D Face Reconstruction: Seeing Through Occlusions 3935
Anh Tuan Tran (Institute for Robotics and Intelligent Systems, USC, CA, USA), Tal Hassner (Information Sciences Institute, USC, CA, USA; The Open University of Israel, Israel), Iacopo Masi (Institute for Robotics and Intelligent Systems, USC, CA, USA), Eran Paz (The Open University of Israel, Israel), Yuval Nirkin (The Open University of Israel, Israel), and Gérard Medioni (Institute for Robotics and Intelligent Systems, USC, CA, USA)

Beyond Grobner Bases: Basis Selection for Minimal Solvers 3945
Viktor Larsson (Lund University), Magnus Oskarsson (Lund University), Kalle Astrom (Lund University), Alge Wallis (University of Bath Bath, United Kingdom), Tomas Pajdla (CIIRC Czech Technical University in Prague, Czech Republic), and Zuzana Kukelova (FEE Czech Technical University in Prague, Czech Republic)

Lions and Tigers and Bears: Capturing Non-rigid, 3D, Articulated Shape from Images 3955
Silvia Zuffi (IMATI-CNR, Milan, Italy), Angjoo Kanazawa (University of California, Berkeley), and Michael J. Black (Max Planck Institute for Intelligent Systems, Tübingen, Germany)

Poster Session P2-1

Deep Cocktail Network: Multi-source Unsupervised Domain Adaptation with Category Shift 3964
Ruijia Xu (Sun Yat-sen University), Ziliang Chen (Sun Yat-sen University), Wangmeng Zuo (Harbin Institute of Technology), Junjie Yan (SenseTime Research), and Liang Lin (Sun Yat-sen University; SenseTime Research)

DOTA: A Large-Scale Dataset for Object Detection in Aerial Images 3974
Gui-Song Xia (Wuhan University), Xiang Bai (Huazhong University Sci. and Tech.), Jian Ding (Wuhan University), Zhen Zhu (Huazhong University Sci. and Tech.), Serge Belongie (Cornell University), Jiebo Luo (Rochester University), Mihai Datcu (German Aerospace Center (DLR)), Marcello Pelillo (University of Venice), and Liangpei Zhang (Wuhan University)

Finding Beans in Burgers: Deep Semantic-Visual Embedding with Localization 3984
Martin Engilberge (Sorbonne Université, Paris, France; Technicolor, Cesson Sévigné, France), Louis Chevallier (Technicolor, Cesson Sévigné, France), Patrick Perez (Technicolor, Cesson Sévigné, France), and Matthieu Cord (Sorbonne Université, Paris, France)

Feature Super-Resolution: Make Machine See More Clearly 3994
Weimin Tan (Fudan University), Bo Yan (Fudan University), and Bahetiyaer Bare (Fudan University)

ClusterNet: Detecting Small Objects in Large Scenes by Exploiting Spatio-Temporal Information 4003
Rodney LaLonde (University of Central Florida), Dong Zhang (University of Central Florida), and Mubarak Shah (University of Central Florida)

MaskLab: Instance Segmentation by Refining Object Detection with Semantic and Direction Features 4013
Liang-Chieh Chen (Google Inc.), Alexander Hermans (RWTH Aachen University), George Papandreou (Google Inc.), Florian Schroff (Google Inc.), Peng Wang (UCLA), and Hartwig Adam (Google Inc.)

Hashing as Tie-Aware Learning to Rank 4023
Kun He (Boston University), Fatih Cakir (Boston University), Sarah Adel Bargal (Boston University), and Stan Sclaroff (Boston University)

Classification-Driven Dynamic Image Enhancement 4033
Vivek Sharma (CV:HCI, KIT, Karlsruhe), Ali Diba (ESAT-PSI, KU Leuven), Davy Neven (ESAT-PSI, KU Leuven), Michael S. Brown (York University, Toronto), Luc Van Gool (ESAT-PSI, KU Leuven; CVL, ETH Zurich), and Rainer Stiefelhagen (CV:HCI, KIT, Karlsruhe)

Knowledge Aided Consistency for Weakly Supervised Phrase Grounding 4042
Kan Chen (University of Southern California, Los Angeles, USA), Jiyang Gao (University of Southern California, Los Angeles, USA), and Ram Nevatia (University of Southern California, Los Angeles, USA)

Who Let the Dogs Out? Modeling Dog Behavior from Visual Data 4051
Kiana Ehsani (University of Washington), Hessam Bagherinezhad (University of Washington), Joseph Redmon (University of Washington), Roozbeh Mottaghi (Allen Institute for AI (AI2)), and Ali Farhadi (University of Washington; Allen Institute for AI (AI2))

Pseudo Mask Augmented Object Detection 4061
Xiangyun Zhao (Northwestern University), Shuang Liang (Tongji University), and Yichen Wei (Microsoft Research)

Dual Skipping Networks 4071
Changmao Cheng (Fudan University, China), Yanwei Fu (Fudan University, China), Yu-Gang Jiang (Fudan University, China), Wei Liu (Tencent AI Lab), Wenlian Lu (Fudan University, China), Jianfeng Feng (Fudan University, China), and Xiangyang Xue (Fudan University, China)

Memory Matching Networks for One-Shot Image Recognition 4080
Qi Cai (University of Science and Technology of China, Hefei, China), Yingwei Pan (University of Science and Technology of China, Hefei, China), Ting Yao (Microsoft Research, Beijing, China), Chenggang Yan (Hangzhou Dianzi University, Hangzhou, China), and Tao Mei (Microsoft Research, Beijing, China)

IQA: Visual Question Answering in Interactive Environments 4089
Daniel Gordon (University of Washington), Aniruddha Kembhavi (Allen Institute for Artificial Intelligence), Mohammad Rastegari (Allen Institute for Artificial Intelligence; Xnor.ai), Joseph Redmon (University of Washington), Dieter Fox (University of Washington; Nvidia), and Ali Farhadi (University of Washington; Allen Institute for Artificial Intelligence)

Pose Transferrable Person Re-identification 4099
Jinxian Liu (Shanghai Jiao Tong University, Shanghai, China), Bingbing Ni (Shanghai Jiao Tong University, Shanghai, China), Yichao Yan (Shanghai Jiao Tong University, Shanghai, China), Peng Zhou (Shanghai Jiao Tong University, Shanghai, China), Shuo Cheng (Shanghai Jiao Tong University, Shanghai, China), and Jianguo Hu (Minivision)

Large Scale Fine-Grained Categorization and Domain-Specific Transfer Learning 4109
Yin Cui (Cornell University; Comell Tech), Yang Song (Google Research), Chen Sun (Google Research), Andrew Howard (Google Research), and Serge Belongie (Cornell University; Comell Tech)

Data Distillation: Towards Omni-Supervised Learning 4119
Ilija Radosavovic (Facebook AI Research (FAIR)), Piotr Dollár (Facebook AI Research (FAIR)), Ross Girshick (Facebook AI Research (FAIR)), Georgia Gkioxari (Facebook AI Research (FAIR)), and Kaiming He (Facebook AI Research (FAIR))

Object Referring in Videos with Language and Human Gaze 4129
Arun Balajee Vasudevan (ETH Zurich), Dengxin Dai (ETH Zurich), and Luc Van Gool (ETH Zurich; KU Leuven)

Feature Selective Networks for Object Detection 4139
Yao Zhai (University of Science and Technology of China), Jingjing Fu (Microsoft Research Asia), Yan Lu (Microsoft Research Asia), and Houqiang Li (University of Science and Technology of China)

Learning a Discriminative Filter Bank Within a CNN for Fine-Grained Recognition 4148
Yaming Wang (University of Maryland, College Park), Vlad I. Morariu (Adobe Research), and Larry S. Davis (University of Maryland, College Park)

Grounding Referring Expressions in Images by Variational Context 4158
Hanwang Zhang (Nanyang Technological University), Yulei Niu (Renmin University of China), and Shih-Fu Chang (Columbia University)

Dynamic Graph Generation Network: Generating Relational Knowledge from Diagrams 4167
Daesik Kim (Seoul National University; V.DO Inc.), YoungJoon Yoo (Clova AI Research, NAVER Corp.), Jeesoo Kim (Seoul National University), Sangkuk Lee (Seoul National University; V.DO Inc.), and Nojun Kwak (Seoul National University)

A Network Architecture for Point Cloud Classification via Automatic Depth Images Generation 4176
Riccardo Roveri (ETH Zurich), Lukas Rahmann (ETH Zurich), A. Cengiz Öztireli (Disney Research Zurich), and Markus Gross (ETH Zurich)

Towards Dense Object Tracking in a 2D Honeybee Hive 4185
Katarzyna Bozek (OIST Graduate University Okinawa, Japan), Laetitia Hebert (OIST Graduate University Okinawa, Japan), Alexander S. Mikheyev (OIST Graduate University Okinawa, Japan), and Greg J. Stephens (OIST Graduate University Okinawa, Japan; VU University, The Netherlands)

Long-Term On-board Prediction of People in Traffic Scenes Under Uncertainty 4194
Apratim Bhattacharyya (Max Planck Institute for Informatics, Saarland Informatics Campus, Saarbrücken, Germany), Mario Fritz (Max Planck Institute for Informatics, Saarland Informatics Campus, Saarbrücken, Germany), and Bernt Schiele (Max Planck Institute for Informatics, Saarland Informatics Campus, Saarbrücken, Germany)

Single-Shot Refinement Neural Network for Object Detection 4203
Shifeng Zhang (CBSR & NLPR, Institute of Automation, Chinese Academy of Sciences, Beijing, China; University of Chinese Academy of Sciences, Beijing, China), Longyin Wen (GE Global Research, Niskayuna, NY), Xiao Bian (GE Global Research, Niskayuna, NY), Zhen Lei (CBSR & NLPR, Institute of Automation, Chinese Academy of Sciences, Beijing, China; University of Chinese Academy of Sciences, Beijing, China), and Stan Z. Li (CBSR & NLPR, Institute of Automation, Chinese Academy of Sciences, Beijing, China; University of Chinese Academy of Sciences, Beijing, China; Macau University of Science and Technology, Macau, China)

Video Captioning via Hierarchical Reinforcement Learning 4213
Xin Wang (University of California, Santa Barbara), Wenhu Chen (University of California, Santa Barbara), Jiawei Wu (University of California, Santa Barbara), Yuan-Fang Wang (University of California, Santa Barbara), and William Yang Wang (University of California, Santa Barbara)

Tips and Tricks for Visual Question Answering: Learnings from the 2017 Challenge 4223
Damien Teney (University of Adelaide), Peter Anderson (Australian National University), Xiaodong He (JD AI Research), and Anton van den Hengel (University of Adelaide)

Learning to Segment Every Thing 4233
Ronghang Hu (BAIR, UC Berkeley; Facebook AI Research (FAIR)), Piotr Dollár (Facebook AI Research (FAIR)), Kaiming He (Facebook AI Research (FAIR)), Trevor Darrell (BAIR, UC Berkeley), and Ross Girshick (Facebook AI Research (FAIR))

Self-Supervised Adversarial Hashing Networks for Cross-Modal Retrieval 4242
Chao Li (Xidian University, Xi'an, China), Cheng Deng (Xidian University, Xi'an, China), Ning Li (Xidian University, Xi'an, China), Wei Liu (Tencent AI Lab, Shenzhen, China), Xinbo Gao (Xidian University, Xi'an, China), and Dacheng Tao (UBTECH Sydney AI Centre, SIT, FEIT, University of Sydney, Australia)

Parallel Attention: A Unified Framework for Visual Object Discovery Through Dialogs and Queries 4252
Bohan Zhuang (University of Adelaide, Australia; Australian Centre for Robotic Vision), Qi Wu (University of Adelaide, Australia; Australian Centre for Robotic Vision), Chunhua Shen (University of Adelaide, Australia; Australian Centre for Robotic Vision), Ian Reid (University of Adelaide, Australia; Australian Centre for Robotic Vision), and Anton van den Hengel (University of Adelaide, Australia; Australian Centre for Robotic Vision)

Zigzag Learning for Weakly Supervised Object Detection 4262
Xiaopeng Zhang (National University of Singapore), Jiashi Feng (National University of Singapore), Hongkai Xiong (Shanghai Jiao Tong University), and Qi Tian (University of Texas at San Antonio)

Attentive Fashion Grammar Network for Fashion Landmark Detection and Clothing Category Classification 4271
Wenguan Wang (Beijing Institute of Technology, China), Wenguan Wang (University of California, Los Angeles, USA), Yuanlu Xu (University of California, Los Angeles, USA), Jianbing Shen (Beijing Institute of Technology, China), and Song-Chun Zhu (University of California, Los Angeles, USA)

Generalized Zero-Shot Learning via Synthesized Examples 4281
Vinay Kumar Verma (Indian Institute of Technology Kanpur), Gundeep Arora (Indian Institute of Technology Kanpur), Ashish Mishra (Indian Institute of Technology Madras), and Piyush Rai (Indian Institute of Technology Kanpur)

Partially Shared Multi-task Convolutional Neural Network with Local Constraint for Face Attribute Learning 4290
Jiajiong Cao (Zhejiang University, China), Yingming Li (Zhejiang University, China), and Zhongfei Zhang (Zhejiang University, China)

SYQ: Learning Symmetric Quantization for Efficient Deep Neural Networks 4300
Julian Faraone (The University of Sydney), Nicholas Fraser (Xilinx Research Labs), Michaela Blott (Xilinx Research Labs), and Philip H.W. Leong (The University of Sydney)

DS*: Tighter Lifting-Free Convex Relaxations for Quadratic Matching Problems 4310
Florian Bernard (MPI Informatics), Florian Bernard (Saarland Informatics Campus), Christian Theobalt (MPI Informatics; Saarland Informatics Campus), and Michael Moeller (University of Siegen)

Deep Mutual Learning 4320
Ying Zhang (Dalian University of Technology, China; Queen Mary University of London), Tao Xiang (Queen Mary University of London), Timothy M. Hospedales (University of Edinburgh, UK), and Huchuan Lu (Dalian University of Technology, China)

Coupled End-to-End Transfer Learning with Generalized Fisher Information 4329
Shixing Chen (Wayne State University), Caojin Zhang (Wayne State University), and Ming Dong (Wayne State University)

Residual Parameter Transfer for Deep Domain Adaptation 4339
Artem Rozantsev (Computer Vision Laboratory, École Polytechnique Fédérale de Lausanne Lausanne, Switzerland), Mathieu Salzmann (Computer Vision Laboratory, École Polytechnique Fédérale de Lausanne Lausanne, Switzerland), and Pascal Fua (Computer Vision Laboratory, École Polytechnique Fédérale de Lausanne Lausanne, Switzerland)

High-Order Tensor Regularization with Application to Attribute Ranking 4349
Kwang In Kim (University of Bath), Juhyun Park (Lancaster University), and James Tompkin (Brown University)

Learning to Localize Sound Source in Visual Scenes 4358
Arda Senocak (KAIST, South Korea), Tae-Hyun Oh (MIT CSAIL, MA, USA), Junsik Kim (KAIST, South Korea), Ming-Hsuan Yang (University of California, Merced, CA, USA), and In So Kweon (KAIST, South Korea)

Dynamic Few-Shot Visual Learning Without Forgetting 4367
Spyros Gidaris (University Paris-Est, LIGM Ecole des Ponts ParisTech) and Nikos Komodakis (University Paris-Est, LIGM Ecole des Ponts ParisTech)

Two-Step Quantization for Low-bit Neural Networks 4376
Peisong Wang (Institute of Automation, Chinese Academy of Sciences, Beijing, China; University of Chinese Academy of Sciences, Beijing, China), Qinghao Hu (Institute of Automation, Chinese Academy of Sciences, Beijing, China; University of Chinese Academy of Sciences, Beijing, China), Yifan Zhang (Institute of Automation, Chinese Academy of Sciences, Beijing, China; University of Chinese Academy of Sciences, Beijing, China), Chunjie Zhang (Institute of Automation, Chinese Academy of Sciences, Beijing, China; University of Chinese Academy of Sciences, Beijing, China), Yang Liu (Alibaba Group, Hangzhou, China), and Jian Cheng (Institute of Automation, Chinese Academy of Sciences, Beijing, China; University of Chinese Academy of Sciences, Beijing, China; Center for Excellence in Brain Science and Intelligence Technology, CAS, Beijing, China)

Improved Lossy Image Compression with Priming and Spatially Adaptive Bit Rates for Recurrent Networks 4385
Nick Johnston (Google Research), Damien Vincent (Google Research), David Minnen (Google Research), Michele Covell (Google Research), Saurabh Singh (Google Research), Troy Chinen (Google Research), Sung Jin Hwang (Google Research), Joel Shor (Google Research), and George Toderici (Google Research)

Conditional Probability Models for Deep Image Compression 4394
Fabian Mentzer (ETH Zurich, Switzerland), Eirikur Agustsson (ETH Zurich, Switzerland), Michael Tschannen (ETH Zurich, Switzerland), Radu Timofte (ETH Zurich, Switzerland), and Luc Van Gool (ETH Zurich, Switzerland)

Deep Diffeomorphic Transformer Networks 4403
Nicki Skafte Detlefsen (Technical University of Denmark), Oren Freifeld (Ben-Gurion University), and Søren Hauberg (Technical University of Denmark)

The Lovasz-Softmax Loss: A Tractable Surrogate for the Optimization of the Intersection-Over-Union Measure in Neural Networks 4413
Maxim Berman (Center for Processing Speech and Images KU Leuven, Belgium), Amal Rannen Triki (Center for Processing Speech and Images KU Leuven, Belgium), and Matthew B. Blaschko (Center for Processing Speech and Images KU Leuven, Belgium)

Generative Adversarial Perturbations 4422
Omid Poursaeed (Cornell University; Cornell Tech), Isay Katsman (Cornell University), Bicheng Gao (Shanghai Jiao Tong University; Cornell University), and Serge Belongie (Cornell University; Cornell Tech)

Learning Strict Identity Mappings in Deep Residual Networks 4432
Xin Yu (University of Utah), Zhiding Yu (NVIDIA), and Srikumar Ramalingam (University of Utah)

Geometric Robustness of Deep Networks: Analysis and Improvement 4441
Can Kanbak (École Polytechnique Fédérale de Lausanne), Seyed-Mohsen Moosavi-Dezfooli (École Polytechnique Fédérale de Lausanne), and Pascal Frossard (École Polytechnique Fédérale de Lausanne)

View Extrapolation of Human Body from a Single Image 4450
Hao Zhu (Nanjing University, Nanjing, China), Hao Zhu (University of Kentucky, Lexington, KY, USA), Hao Su (University of California, San Diego, CA, USA), Peng Wang (Baidu Inc., Beijing, China; National Engineering Laboratory of Deep Learning and Technology and Application, China), Xun Cao (Nanjing University, Nanjing, China), and Ruigang Yang (University of Kentucky, Lexington, KY, USA; Baidu Inc., Beijing, China; National Engineering Laboratory of Deep Learning and Technology and Application, China)

Geometry Aware Constrained Optimization Techniques for Deep Learning 4460
Soumava Kumar Roy (Australian National University, Canberra, Australia), Zakaria Mhammedi (Australian National University, Canberra, Australia; Data61, CSIRO, Canberra, Australia), and Mehrtash Harandi (Australian National University, Canberra, Australia; Data61, CSIRO, Canberra, Australia)

PointNetVLAD: Deep Point Cloud Based Retrieval for Large-Scale Place Recognition 4470
Mikaela Angelina Uy (National University of Singapore) and Gim Hee Lee (National University of Singapore)

An Efficient and Provable Approach for Mixture Proportion Estimation Using Linear Independence Assumption 4480
Xiyu Yu (The University of Sydney, Australia), Tongliang Liu (The University of Sydney, Australia), Mingming Gong (University of Pittsburgh; Carnegie Mellon University), Kayhan Batmanghelich (University of Pittsburgh), and Dacheng Tao (The University of Sydney, Australia)

VoxelNet: End-to-End Learning for Point Cloud Based 3D Object Detection 4490
Yin Zhou (Apple Inc) and Oncel Tuzel (Apple Inc)

Image to Image Translation for Domain Adaptation 4500
Zak Murez (University of California, San Diego), Zak Murez (HRL Laboratories, LLC), Soheil Kolouri (HRL Laboratories, LLC), David Kriegman (University of California, San Diego), Ravi Ramamoorthi (University of California, San Diego), and Kyungnam Kim (HRL Laboratories, LLC)

MobileNetV2: Inverted Residuals and Linear Bottlenecks 4510
Mark Sandler (Google Inc.), Andrew Howard (Google Inc.), Menglong Zhu (Google Inc.), Andrey Zhmoginov (Google Inc.), and Liang-Chieh Chen (Google Inc.)

Im2Struct: Recovering 3D Shape Structure from a Single RGB Image 4521
Chengjie Niu (National University of Defense Technology), Jun Li (National University of Defense Technology), and Kai Xu (National University of Defense Technology)

Trust your Model: Light Field Depth Estimation with Inline Occlusion Handling 4530
Hendrik Schilling (Heidelberg Collaboratory for Image Processing (HCI)), Maximilian Diebold (Heidelberg Collaboratory for Image Processing (HCI)), Carsten Rother (Heidelberg Collaboratory for Image Processing (HCI)), and Bernd Jähne (Heidelberg Collaboratory for Image Processing (HCI))

Baseline Desensitizing in Translation Averaging 4539
Bingbing Zhuang (National University of Singapore), Loong-Fah Cheong (National University of Singapore), and Gim Hee Lee (National University of Singapore)

Mining Point Cloud Local Structures by Kernel Correlation and Graph Pooling 4548
Yiru Shen (Clemson University), Chen Feng (Mitsubishi Electric Research Laboratories (MERL)), Yaoqing Yang (Carnegie Mellon University), and Dong Tian (Mitsubishi Electric Research Laboratories (MERL))

Large-Scale Point Cloud Semantic Segmentation with Superpoint Graphs 4558
Loic Landrieu (Universite Paris-Est, LASTIG MATIS IGN, ENSG) and Martin Simonovsky (Universite Paris-Est, Ecole des Ponts ParisTech)

Very Large-Scale Global SfM by Distributed Motion Averaging 4568
Siyu Zhu (The Hong Kong University of Science and Technology), Runze Zhang (The Hong Kong University of Science and Technology), Lei Zhou (The Hong Kong University of Science and Technology), Tianwei Shen (The Hong Kong University of Science and Technology), Tian Fang (Altizure.com), Ping Tan (Simon Fraser University), and Long Quan (The Hong Kong University of Science and Technology)

ScanComplete: Large-Scale Scene Completion and Semantic Segmentation for 3D Scans 4578
Angela Dai (Stanford University; Google; DeepMind), Daniel Ritchie (Brown University), Martin Bokeloh (Google), Scott Reed (DeepMind), Jürgen Sturm (Google), and Matthias Nießner (Technical University of Munich)

Solving the Perspective-2-Point Problem for Flying-Camera Photo Composition 4588
Ziquan Lan (NUS Graduate School for Integrative Sciences and Engineering; National University of Singapore), David Hsu (NUS Graduate School for Integrative Sciences and Engineering; National University of Singapore), and Gim Hee Lee

Reflection Removal for Large-Scale 3D Point Clouds 4597
Jae-Seong Yun (Ulsan National Institute of Science and Technology, Ulsan, Korea) and Jae-Young Sim (Ulsan National Institute of Science and Technology, Ulsan, Korea)

Attentional ShapeContextNet for Point Cloud Recognition 4606
Saining Xie (University of California, San Diego), Sainan Liu (University of California, San Diego), Zeyu Chen (University of California, San Diego), and Zhuowen Tu (University of California, San Diego)

Geometry-Aware Deep Network for Single-Image Novel View Synthesis 4616
Miaomiao Liu (CECS, ANU Canberra, Australia), Xuming He (ShanghaiTech University Shanghai, China), and Mathieu Salzmann (CVLAB, EPFL Lausanne, Switzerland)

InverseFaceNet: Deep Monocular Inverse Face Rendering 4625
Hyeongwoo Kim (Max-Planck-Institute for Informatics; Saarland Informatics Campus), Justus Thies (Technical University of Munich), Michael Zollhöfer (Max-Planck-Institute for Informatics; Saarland Informatics Campus; Stanford University), Christian Richardt (University of Bath), Ayush Tewari (Max-Planck-Institute for Informatics; Saarland Informatics Campus), and Christian Theobalt (Max-Planck-Institute for Informatics; Saarland Informatics Campus)

Sparse Photometric 3D Face Reconstruction Guided by Morphable Models 4635
Xuan Cao (ShanghaiTech University, Shanghai, China), Zhang Chen (ShanghaiTech University, Shanghai, China), Anpei Chen (ShanghaiTech University, Shanghai, China), Xin Chen (ShanghaiTech University, Shanghai, China), Shiying Li (ShanghaiTech University, Shanghai, China), and Jingyi Yu (ShanghaiTech University, Shanghai, China)

Texture Mapping for 3D Reconstruction with RGB-D Sensor 4645
Yanping Fu (Wuhan University, China), Qingan Yan (JD.com), Long Yang (Northwest A&F University, China), Jie Liao (Wuhan University, China), and Chunxia Xiao (Wuhan University, China)

Learning Less is More - 6D Camera Localization via 3D Surface Regression 4654
Eric Brachmann (Visual Learning Lab Heidelberg University (HCI/IWR)) and Carsten Rother (Visual Learning Lab Heidelberg University (HCI/IWR))

Feature Mapping for Learning Fast and Accurate 3D Pose Inference from Synthetic Images 4663
Mahdi Rad (Institute for Computer Graphics and Vision, Graz University of Technology, Graz, Austria), Markus Oberweger (Institute for Computer Graphics and Vision, Graz University of Technology, Graz, Austria), and Vincent Lepetit (Institute for Computer Graphics and Vision, Graz University of Technology, Graz, Austria; Laboratoire Bordelais de Recherche en Informatique, UniversitÃ© de Bordeaux, Bordeaux, France)

Indoor RGB-D Compass from a Single Line and Plane 4673
Pyojin Kim (ASRI, Seoul National University, South Korea), Brian Coltin (SGT, Inc., NASA Ames Research Center, USA), and H. Jin Kim (ASRI, Seoul National University, South Korea)

Geometry-Aware Network for Non-rigid Shape Prediction from a Single View 4681
A. Pumarola (Institut de Robòtica i Informàtica Industrial, CSIC-UPC, Barcelona, Spain), A. Agudo (Institut de Robòtica i Informàtica Industrial, CSIC-UPC, Barcelona, Spain), L. Porzi (Mapillary Research, Graz, Austria), A. Sanfeliu (Institut de Robòtica i Informàtica Industrial, CSIC-UPC, Barcelona, Spain), V. Lepetit (Laboratoire Bordelais de Recherche en Informatique, UniversitÃ© de Bordeaux, France), and F. Moreno-Noguer (Institut de Robòtica i Informàtica Industrial, CSIC-UPC, Barcelona, Spain)

Sim2Real Viewpoint Invariant Visual Servoing by Recurrent Control 4691
Fereshteh Sadeghi (University of Washington; Google Brain), Alexander Toshev (Google Brain), Eric Jang (Google Brain), and Sergey Levine (University of California, Berkeley; Google Brain)

DocUNet: Document Image Unwarping via a Stacked U-Net 4700
Ke Ma (Stony Brook University), Zhixin Shu (Stony Brook University), Xue Bai (Megvii Inc.), Jue Wang (Megvii Inc.), and Dimitris Samaras (Stony Brook University)

Analysis of Hand Segmentation in the Wild 4710
Aisha Urooj Khan (Center for Research in Computer Vision, University of Central Florida) and Ali Borji (Center for Research in Computer Vision, University of Central Florida)

RoadTracer: Automatic Extraction of Road Networks from Aerial Images 4720
Favyen Bastani (MIT CSAIL), Songtao He (MIT CSAIL), Sofiane Abbar (Qatar Computing Research Institute, HBKU), Mohammad Alizadeh (MIT CSAIL), Hari Balakrishnan (MIT CSAIL), Sanjay Chawla (Qatar Computing Research Institute, HBKU), Sam Madden (MIT CSAIL), and David DeWitt (MIT CSAIL)

Alternating-Stereo VINS: Observability Analysis and Performance Evaluation 4729
Mrinal K. Paul (Google Inc.) and Stergios I. Roumeliotis (Google Inc.)

Soccer on Your Tabletop 4738
Konstantinos Rematas (University of Washington), Ira Kemelmacher-Shlizerman (University of Washington; Facebook), Brian Curless (University of Washington), and Steve Seitz (University of Washington; Google)

Poster Session P2-2

EPINET: A Fully-Convolutional Neural Network Using Epipolar Geometry for Depth from Light Field Images 4748
Changha Shin (Yonsei University), Hae-Gon Jeon (KAIST), Youngjin Yoon (KAIST), In So Kweon (KAIST), and Seon Joo Kim (Yonsei University)

A Hybrid l1-l0 Layer Decomposition Model for Tone Mapping 4758
Zhetong Liang (The Hong Kong Polytechnic University), Jun Xu (The Hong Kong Polytechnic University), David Zhang (The Hong Kong Polytechnic University), Zisheng Cao (DJI Co.), and Lei Zhang (The Hong Kong Polytechnic University)

Deeply Learned Filter Response Functions for Hyperspectral Reconstruction 4767
Shijie Nie (National Institute of Informatics), Lin Gu (National Institute of Informatics), Yinqiang Zheng (National Institute of Informatics), Antony Lam (Saitama University), Nobutaka Ono (Tokyo Metropolitan University), and Imari Sato (National Institute of Informatics)

CRRN: Multi-scale Guided Concurrent Reflection Removal Network 4777
Renjie Wan (Nanyang Technological University, Singapore), Boxin Shi (Peking University, China), Ling-Yu Duan (Peking University, China), Ah-Hwee Tan (Nanyang Technological University, Singapore), and Alex C. Kot (Nanyang Technological University, Singapore)

Single Image Reflection Separation with Perceptual Losses 4786
Xuaner Zhang (UC Berkeley), Ren Ng (UC Berkeley), and Qifeng Chen (Intel Labs)

A Robust Method for Strong Rolling Shutter Effects Correction Using Lines with Automatic Feature Selection 4795
Yizhen Lao (Institut Pascal, Universite Clermont Auvergne, France) and Omar Ait-Aider (Institut Pascal, Universite Clermont Auvergne, France)

Time-Resolved Light Transport Decomposition for Thermal Photometric Stereo 4804
Kenichiro Tanaka (Nara Institute of Science and Technology (NAIST), Japan), Hiroyuki Kubo (Nara Institute of Science and Technology (NAIST), Japan), Nobuhiro Ikeya (Nara Institute of Science and Technology (NAIST), Japan), Takuya Funatomi (Nara Institute of Science and Technology (NAIST), Japan), Tsuyoshi Takatani (Nara Institute of Science and Technology (NAIST), Japan), and Yasuhiro Mukaigawa (Nara Institute of Science and Technology (NAIST), Japan)

Efficient Diverse Ensemble for Discriminative Co-tracking 4814
Kourosh Meshgi (Kyoto University, Kyoto, Japan), Shigeyuki Oba (Kyoto University, Kyoto, Japan), and Shin Ishii (Kyoto University, Kyoto, Japan)

Rolling Shutter and Radial Distortion are Features for High Frame Rate Multi-camera Tracking 4824
Akash Bapat (The University of North Carolina at Chapel Hill), True Price (The University of North Carolina at Chapel Hill), and Jan-Michael Frahm (The University of North Carolina at Chapel Hill)

A Twofold Siamese Network for Real-Time Object Tracking 4834
Anfeng He (University of Science and Technology of China, Hefei, Anhui, China), Chong Luo (Microsoft Research, Beijing, China), Xinmei Tian (University of Science and Technology of China, Hefei, Anhui, China), and Wenjun Zeng (Microsoft Research, Beijing, China)

Multi-cue Correlation Filters for Robust Visual Tracking 4844
Ning Wang (CAS Key Laboratory of GIPAS, University of Science and Technology of China), Wengang Zhou (CAS Key Laboratory of GIPAS, University of Science and Technology of China), Qi Tian (University of Texas at San Antonio), Richang Hong (HeFei University of Technology), Meng Wang (HeFei University of Technology), and Houqiang Li (CAS Key Laboratory of GIPAS, University of Science and Technology of China)

Learning Attentions: Residual Attentional Siamese Network for High Performance Online Visual Tracking 4854
Qiang Wang (University of Chinese Academy of Sciences, Beijing, China; National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China), Zhu Teng (School of Computer and Information Technology, Beijing Jiaotong University, Beijing, China), Junliang Xing (National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China), Jin Gao (National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China), Weiming Hu (University of Chinese Academy of Sciences, Beijing, China; National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China), and Stephen Maybank (Birkbeck College, University of London)

SINT++: Robust Visual Tracking via Adversarial Positive Instance Generation 4864
Xiao Wang (Anhui University, Hefei, China), Chenglong Li (Anhui University, Hefei, China), Bin Luo (Anhui University, Hefei, China), and Jin Tang (Anhui University, Hefei, China)

High-Speed Tracking with Multi-kernel Correlation Filters 4874
Ming Tang (National Lab of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China), Bin Yu (National Lab of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China), Fan Zhang (School of Info. & Comm. Eng., Beijing University of Posts and Telecommunications), and Jinqiao Wang (National Lab of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China)

Occlusion Aware Unsupervised Learning of Optical Flow 4884
Yang Wang (Baidu Research), Yi Yang (Baidu Research), Zhenheng Yang (University of Southern California), Liang Zhao (Baidu Research), Peng Wang (Baidu Research), and Wei Xu (Baidu Research; National Engineering Laboratory for Deep Learning Technology and Applications)

Revisiting Video Saliency: A Large-Scale Benchmark and a New Model 4894
Wenguan Wang (Beijing Institute of Technology, China), Jianbing Shen (Beijing Institute of Technology, China), Fang Guo (Beijing Institute of Technology, China), Ming-Ming Cheng (Nankai University, China), and Ali Borji (University of Central Florida, USA)

Learning Spatial-Temporal Regularized Correlation Filters for Visual Tracking 4904
Feng Li (Harbin Institute of Technology, China), Cheng Tian (Harbin Institute of Technology, China), Wangmeng Zuo (Harbin Institute of Technology, China), Lei Zhang (The Hong Kong Polytechnic University, China), and Ming-Hsuan Yang (University of California, Merced, USA)

Multimodal Visual Concept Learning with Weakly Supervised Techniques 4914
Giorgos Bouritsas (National Technical University of Athens, Greece), Petros Koutras (National Technical University of Athens, Greece), Athanasia Zlatintsi (National Technical University of Athens, Greece), and Petros Maragos (National Technical University of Athens, Greece)

Efficient Large-Scale Approximate Nearest Neighbor Search on OpenCL FPGA 4924
Jialiang Zhang (UW-Madison), Jing Li (UW-Madison), and Soroosh Khoram (UW-Madison)

Learning a Complete Image Indexing Pipeline 4933
Himalaya Jain (Technicolor, Rennes, France; Inria, Rennes, France), Joaquin Zepeda (Amazon, Seattle, WA), Patrick Pérez (Technicolor, Rennes, France; Valeo.ai, Paris, France), and Rémi Gribonval (Inria, Rennes, France)

Transparency by Design: Closing the Gap Between Performance and Interpretability in Visual Reasoning..... 4942
David Mascharka (MIT Lincoln Laboratory), Philip Tran (Planck Aerosystems), Ryan Soklaski (MIT Lincoln Laboratory), and Arjun Majumdar (MIT Lincoln Laboratory)

Fooling Vision and Language Models Despite Localization and Attention Mechanism 4951
Xiaojun Xu (Shanghai Jiao Tong University; UC Berkeley), Xinyun Chen (UC Berkeley), Chang Liu (UC Berkeley), Anna Rohrbach (UC Berkeley; MPI for Informatics), Trevor Darrell (UC Berkeley), and Dawn Song (UC Berkeley)

Categorizing Concepts with Basic Level for Vision-to-Language 4962
Hanzhang Wang (Tongji University, Shanghai, P. R. China), Hanli Wang (Tongji University, Shanghai, P. R. China), and Kaisheng Xu (Tongji University, Shanghai, P. R. China)

Don't Just Assume; Look and Answer: Overcoming Priors for Visual Question Answering 4971
Aishwarya Agrawal (Georgia Institute of Technology), Dhruv Batra (Georgia Institute of Technology; Facebook AI Research), Devi Parikh (Georgia Institute of Technology; Facebook AI Research), and Aniruddha Kembhavi (Allen Institute for Artificial Intelligence)

Learning Pixel-Level Semantic Affinity with Image-Level Supervision for Weakly Supervised Semantic Segmentation 4981
Jiwoon Ahn (DGIST, Korea) and Suha Kwak (POSTECH, Korea)

From Lifestyle Vlogs to Everyday Interactions 4991
David F. Fouhey (University of California, Berkeley), Wei-cheng Kuo (University of California, Berkeley), Alexei A. Efros (University of California, Berkeley), and Jitendra Malik (University of California, Berkeley)

Cross-Domain Weakly-Supervised Object Detection Through Progressive Domain Adaptation 5001
Naoto Inoue (The University of Tokyo, Japan), Ryosuke Furuta (The University of Tokyo, Japan), Toshihiko Yamasaki (The University of Tokyo, Japan), and Kiyoharu Aizawa (The University of Tokyo, Japan)

RotationNet: Joint Object Categorization and Pose Estimation Using Multiviews from Unsupervised Viewpoints 5010
Asako Kanezaki (National Institute of Advanced Industrial Science and Technology (AIST)), Yasuyuki Matsushita (Graduate School of Information Science and Technology, Osaka University), and Yoshifumi Nishida (National Institute of Advanced Industrial Science and Technology (AIST))

An End-to-End TextSpotter with Explicit Alignment and Attention 5020
Tong He (University of Adelaide, Australia), Zhi Tian (University of Adelaide, Australia; Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China), Weilin Huang (Malong Technologies Co., Ltd, Shenzhen, China), Chunhua Shen (University of Adelaide, Australia), Yu Qiao (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China), and Changming Sun (CSIRO Data61, Australia)

WILDTRACK: A Multi-camera HD Dataset for Dense Unscripted Pedestrian Detection 5030
Tatjana Chavdarova (Machine Learning group, Idiap Research Institute & École Polytechnique Fédérale de Lausanne), Pierre Baqué (CVLab, École Polytechnique Fédérale de Lausanne), Stéphane Bouquet (CVLab, École Polytechnique Fédérale de Lausanne), Andrii Maksai (CVLab, École Polytechnique Fédérale de Lausanne), Cijo Jose (Machine Learning group, Idiap Research Institute & École Polytechnique Fédérale de Lausanne), Timur Bagautdinov (CVLab, École Polytechnique Fédérale de Lausanne), Louis Lettry (Computer Vision Lab, ETH Zurich), Pascal Fua (CVLab, École Polytechnique Fédérale de Lausanne), Luc Van Gool (Computer Vision Lab, ETH Zurich), and François Fleuret (Machine Learning group, Idiap Research Institute & École Polytechnique Fédérale de Lausanne)

Direct Shape Regression Networks for End-to-End Face Alignment 5040
Xin Miao (University of Texas at Arlington, TX, USA), Xiantong Zhen (Beihang University, Beijing, China; University of Texas at Arlington, TX, USA), Xianglong Liu (Beihang University, Beijing, China), Cheng Deng (Xidian University, Xi'an, China), Vassilis Athitsos (University of Texas at Arlington, TX, USA), and Heng Huang (University of Pittsburgh, PA, USA; University of Texas at Arlington, TX, USA)

Natural and Effective Obfuscation by Head Inpainting 5050
Qianru Sun (Max Planck Institute for Informatics, Saarland Informatics Campus), Liqian Ma (KU-Leuven/PSI, Toyota Motor Europe (TRACE)), Seong Joon Oh (Max Planck Institute for Informatics, Saarland Informatics Campus), Luc Van Gool (KU-Leuven/PSI, Toyota Motor Europe (TRACE); ETH Zurich), Bernt Schiele (Max Planck Institute for Informatics, Saarland Informatics Campus), and Mario Fritz (Max Planck Institute for Informatics, Saarland Informatics Campus)

3D Semantic Trajectory Reconstruction from 3D Pixel Continuum 5060
Jae Shin Yoon (University of Minnesota), Ziwei Li (University of Minnesota), and Hyun Soo Park (University of Minnesota)

Optimizing Filter Size in Convolutional Neural Networks for Facial Action Unit Recognition 5070
Shizhong Han (University of South Carolina, Columbia, SC), Zibo Meng (University of South Carolina, Columbia, SC), Zhiyuan Li (University of South Carolina, Columbia, SC), James O'Reilly (University of South Carolina, Columbia, SC), Jie Cai (University of South Carolina, Columbia, SC), Xiaofeng Wang (University of South Carolina, Columbia, SC), and Yan Tong (University of South Carolina, Columbia, SC)

V2V-PoseNet: Voxel-to-Voxel Prediction Network for Accurate 3D Hand and Human Pose Estimation from a Single Depth Map 5079
Ju Yong Chang (Kwangwoon University), Gyeongsik Moon (ASRI, Seoul National University), and Kyoung Mu Lee (ASRI, Seoul National University)

Ring Loss: Convex Feature Normalization for Face Recognition 5089
Yutong Zheng (Carnegie Mellon University), Dipan K. Pal (Carnegie Mellon University), and Marios Savvides (Carnegie Mellon University)

Adversarially Occluded Samples for Person Re-identification 5098
Houjing Huang (CRIPAC & NLPR, CASIA; University of Chinese Academy of Sciences), Dangwei Li (CRIPAC & NLPR, CASIA; University of Chinese Academy of Sciences), Zhang Zhang (CRIPAC & NLPR, CASIA; University of Chinese Academy of Sciences), Xiaotang Chen (CRIPAC & NLPR, CASIA; University of Chinese Academy of Sciences), Kaiqi Huang (CRIPAC & NLPR, CASIA), and Kaiqi Huang (University of Chinese Academy of Sciences; CAS Center for Excellence in Brain Science and Intelligence Technology)

Classifier Learning with Prior Probabilities for Facial Action Unit Recognition 5108
Yong Zhang (National Laboratory of Pattern Recognition, CASIA; University of Chinese Academy of Sciences), Weiming Dong (National Laboratory of Pattern Recognition, CASIA), Bao-Gang Hu (National Laboratory of Pattern Recognition, CASIA), and Qiang Ji (Rensselaer Polytechnic Institute)

4DFAB: A Large Scale 4D Database for Facial Expression Analysis and Biometric Applications 5117
Shiyang Cheng (Imperial College London), Irene Kotsia (Middlesex University London), Maja Pantic (Imperial College London), and Stefanos Zafeiriou (Imperial College London)

Seeing Small Faces from Robust Anchor's Perspective 5127
Chenchen Zhu (Carnegie Mellon University), Ran Tao (Carnegie Mellon University), Khoa Luu (Carnegie Mellon University), and Marios Savvides (Carnegie Mellon University)

2D/3D Pose Estimation and Action Recognition Using Multitask Deep Learning 5137
Diogo C. Luvizon (Paris Seine University, Cergy, France), David Picard (Paris Seine University, Cergy, France; Sorbonne Université, Paris, France), and Hedi Tabia (Paris Seine University, Cergy, France)

Dense 3D Regression for Hand Pose Estimation 5147
Chengde Wan (ETH Zurich), Thomas Probst (ETH Zurich), Luc Van Gool (ETH Zurich; KU Leuven), and Angela Yao (University of Bonn)

Camera Style Adaptation for Person Re-identification 5157
Zhun Zhong (Xiamen University, China; Centre for Artificial Intelligence, University of Technology Sydney, Australia), Liang Zheng (Centre for Artificial Intelligence, University of Technology Sydney, Australia), Zhedong Zheng (Centre for Artificial Intelligence, University of Technology Sydney, Australia), Shaozi Li (Xiamen University, China), and Yi Yang (Centre for Artificial Intelligence, University of Technology Sydney, Australia)

PoseTrack: A Benchmark for Human Pose Estimation and Tracking 5167
Mykhaylo Andriluka (Google Research), Umar Iqbal (Computer Vision Group, University of Bonn, Germany), Eldar Insafutdinov (MPI for Informatics, Saarbrücken, Germany), Leonid Pishchulin (MPI for Informatics, Saarbrücken, Germany), Anton Milan (Amazon Research), Juergen Gall (Computer Vision Group, University of Bonn, Germany), and Bernt Schiele (MPI for Informatics, Saarbrücken, Germany)

Exploit the Unknown Gradually: One-Shot Video-Based Person Re-identification by Stepwise Learning 5177
Yu Wu (SUSTech-UTS Joint Centre of CIS, Southern China University of Science and Technology; University of Technology Sydney), Yutian Lin (University of Technology Sydney), Xuanyi Dong (University of Technology Sydney), Yan Yan (University of Technology Sydney), Wanli Ouyang (The University of Sydney), and Yi Yang (SUSTech-UTS Joint Centre of CIS, Southern China University of Science and Technology; University of Technology Sydney)

Pose-Robust Face Recognition via Deep Residual Equivariant Mapping 5187
Kaidi Cao (SenseTime Research), Yu Rong (CUHK - SenseTime Joint Lab, The Chinese University of Hong Kong; SenseTime Research), Cheng Li (SenseTime Research), Xiaoou Tang (CUHK - SenseTime Joint Lab, The Chinese University of Hong Kong), and Chen Change Loy (CUHK - SenseTime Joint Lab, The Chinese University of Hong Kong)

DecideNet: Counting Varying Density Crowds Through Attention Guided Detection and Density Estimation.. 5197
Jiang Liu (Carnegie Mellon University), Chenqiang Gao (Chongqing University of Posts and Telecommunications), Deyu Meng (Xi'an Jiaotong University), and Alexander G. Hauptmann (Carnegie Mellon University)

LSTM Pose Machines 5207
Yue Luo (SenseTime Research), Jimmy Ren (SenseTime Research), Zhouxia Wang (SenseTime Research), Wenxiu Sun (SenseTime Research), Jinshan Pan (SenseTime Research), Jianbo Liu (SenseTime Research), Jiahao Pang (SenseTime Research), and Liang Lin (SenseTime Research; Sun Yat-sen University, China)

Disentangling Features in 3D Face Shapes for Joint Face Reconstruction and Recognition 5216
Feng Liu (Sichuan University, China), Ronghang Zhu (Sichuan University, China), Dan Zeng (Sichuan University, China), Qijun Zhao (Sichuan University, China), and Xiaoming Liu (Michigan State University, USA)

Convolutional Sequence to Sequence Model for Human Dynamics 5226
Chen Li (National University of Singapore, Singapore), Zhen Zhang (National University of Singapore, Singapore), Wee Sun Lee (National University of Singapore, Singapore), and Gim Hee Lee (National University of Singapore, Singapore)

Gesture Recognition: Focus on the Hands 5235
Pradyumna Narayana (Colorado State University), J. Ross Beveridge (Colorado State University), and Bruce A. Draper (Colorado State University)

Crowd Counting via Adversarial Cross-Scale Consistency Pursuit 5245
Zan Shen (Shanghai Institute for Advanced Communication and Data Science), Yi Xu (Shanghai Jiao Tong University, Shanghai, China), Bingbing Ni (Shanghai Institute for Advanced Communication and Data Science), Minsi Wang (Shanghai Institute for Advanced Communication and Data Science), Jianguo Hu (Minivision), and Xiaokang Yang (Shanghai Institute for Advanced Communication and Data Science)

3D Human Pose Estimation in the Wild by Adversarial Learning 5255
Wei Yang (CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong), Wanli Ouyang (The University of Sydney), Xiaolong Wang (The Robotics Institute, Carnegie Mellon University), Jimmy Ren (SenseTime Research), Hongsheng Li (CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong), and Xiaogang Wang (CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong)

CosFace: Large Margin Cosine Loss for Deep Face Recognition 5265
Hao Wang (Tencent AI Lab), Yitong Wang (Tencent AI Lab), Zheng Zhou (Tencent AI Lab), Xing Ji (Tencent AI Lab), Dihong Gong (Tencent AI Lab), Jingchao Zhou (Tencent AI Lab), Zhifeng Li (Tencent AI Lab), and Wei Liu (Tencent AI Lab)

Encoding Crowd Interaction with Deep Neural Network for Pedestrian Trajectory Prediction 5275
Yanyu Xu (ShanghaiTech University), Zhixin Piao (ShanghaiTech University), and Shenghua Gao (ShanghaiTech University)

Mean-Variance Loss for Deep Age Estimation from a Face 5285
Hongyu Pan (Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China; University of Chinese Academy of Sciences, Beijing, China), Hu Han (Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China), Shiguang Shan (Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China; University of Chinese Academy of Sciences, Beijing, China; CAS Center for Excellence in Brain Science and Intelligence Technology), and Xilin Chen (Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China; University of Chinese Academy of Sciences, Beijing, China)

Probabilistic Joint Face-Skull Modelling for Facial Reconstruction 5295
Dennis Madsen (University of Basel), Marcel Lüthi (University of Basel), Andreas Schneider (University of Basel), and Thomas Vetter (University of Basel)

Learning Latent Super-Events to Detect Multiple Activities in Videos 5304
AJ Piergiovanni (Indiana University) and Michael S. Ryoo (Indiana University)

Temporal Hallucinating for Action Recognition with Few Still Images 5314
Yali Wang (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China), Lei Zhou (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China; SenseTime Group Limited), and Yu Qiao (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China; The Chinese University of Hong Kong)

Deep Progressive Reinforcement Learning for Skeleton-Based Action Recognition 5323
Yansong Tang (Tsinghua University, China; State Key Lab of Intelligent Technologies and Systems, Tsinghua University, China), Yi Tian (Tsinghua University, China), Jiwen Lu (Tsinghua University, China; State Key Lab of Intelligent Technologies and Systems, Tsinghua University, China; Beijing National Research Center for Information Science and Technology, China), Peiyang Li (Tsinghua University, China), and Jie Zhou (Tsinghua University, China; State Key Lab of Intelligent Technologies and Systems, Tsinghua University, China; Beijing National Research Center for Information Science and Technology, China)

Gaze Prediction in Dynamic 360° Immersive Videos 5333
Yanyu Xu (ShanghaiTech University), Yanbing Dong (ShanghaiTech University), Junru Wu (ShanghaiTech University), Zhengzhong Sun (ShanghaiTech University), Zhiru Shi (ShanghaiTech University), Jingyi Yu (ShanghaiTech University), and Shenghua Gao (ShanghaiTech University)

When will you do what? - Anticipating Temporal Occurrences of Activities 5343
Yazan Abu Farha (University of Bonn, Germany), Alexander Richard (University of Bonn, Germany), and Juergen Gall (University of Bonn, Germany)

Fusing Crowd Density Maps and Visual Object Trackers for People Tracking in Crowd Scenes 5353
Weihong Ren (City University of Hong Kong; State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Di Kang (City University of Hong Kong), Yandong Tang (State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences), and Antoni B. Chan (City University of Hong Kong)

Dual Attention Matching Network for Context-Aware Feature Sequence Based Person Re-identification 5363
Jianlou Si (Beijing University of Posts and Telecommunications, Beijing), Honggang Zhang (Beijing University of Posts and Telecommunications, Beijing), Chun-Guang Li (Beijing University of Posts and Telecommunications, Beijing), Jason Kuen (Nanyang Technological University), Xiangfei Kong (Nanyang Technological University), Alex C. Kot (Nanyang Technological University), and Gang Wang (Alibaba AI Labs, Hangzhou)

Easy Identification from Better Constraints: Multi-shot Person Re-identification from Reference Constraints 5373
Jiahuan Zhou (Northwestern University, USA), Bing Su (Institute of Software, Chinese Academy of Science, China), and Ying Wu (Northwestern University, USA)

Crowd Counting with Deep Negative Correlation Learning 5382
Zenglin Shi (University of Bern, Switzerland), Le Zhang (ADSC, UIUC, Singapore), Yun Liu (Nankai University, China), Xiaofeng Cao (University of Technology Sydney, Australia), Yangdong Ye (Zhengzhou University, China), Ming-Ming Cheng (Nankai University, China), and Guoyan Zheng (University of Bern, Switzerland)

Human Appearance Transfer 5391
Mihai Zanfir (Institute of Mathematics of the Romanian Academy), Alin-Ionut Popa (Institute of Mathematics of the Romanian Academy), Andrei Zanfir (Institute of Mathematics of the Romanian Academy), and Cristian Sminchisescu (Lund University; Institute of Mathematics of the Romanian Academy)

Domain Generalization with Adversarial Feature Learning 5400
Haoliang Li (Nanyang Technological University, Singapore), Sinno Jialin Pan (Nanyang Technological University, Singapore), Shiqi Wang (City University of Hong Kong, China), and Alex C. Kot (Nanyang Technological University, Singapore)

Pyramid Stereo Matching Network 5410
Jia-Ren Chang (National Chiao Tung University, Taiwan) and Yong-Sheng Chen (National Chiao Tung University, Taiwan)

Event-Based Vision Meets Deep Learning on Steering Prediction for Self-Driving Cars 5419
Ana I. Maqueda (Grupo de Tratamiento de Imágenes, Universidad Politécnica de Madrid, Spain), Antonio Loquercio (University of Zurich and ETH Zurich, Switzerland), Guillermo Gallego (University of Zurich and ETH Zurich, Switzerland), Narciso Garcí-a (Grupo de Tratamiento de Imágenes, Universidad Politécnica de Madrid, Spain), and Davide Scaramuzza (University of Zurich and ETH Zurich, Switzerland)

Learning Answer Embeddings for Visual Question Answering 5428
Fei Sha (University of Southern California Los Angeles, CA), Wei-Lun Chao (University of Southern California Los Angeles, CA), and Hexiang Hu (University of Southern California Los Angeles, CA)

Good View Hunting: Learning Photo Composition from Dense View Pairs 5437
Zijun Wei (Stony Brook University), Jianming Zhang (Adobe Research), Xiaohui Shen (Adobe Research), Zhe Lin (Adobe Research), Radomír Mech (Adobe Research), Minh Hoai (Stony Brook University), and Dimitris Samaras (Stony Brook University)

CleanNet: Transfer Learning for Scalable Image Classifier Training with Label Noise 5447
Kuang-Huei Lee (Microsoft AI and Research), Xiaodong He (JD AI Research), Lei Zhang (Microsoft AI and Research), and Linjun Yang (Facebook)

Independently Recurrent Neural Network (IndRNN): Building A Longer and Deeper RNN 5457
Shuai Li (University of Wollongong, Australia), Wanqing Li (University of Wollongong, Australia), Chris Cook (University of Wollongong, Australia), Ce Zhu (University of Electronic Science and Technology of China), and Yanbo Gao (University of Electronic Science and Technology of China)

Mix and Match Networks: Encoder-Decoder Alignment for Zero-Pair Image Translation 5467
Yaxing Wang (Computer Vision Center, Universitat Autònoma de Barcelona Barcelona, Spain), Joost van de Weijer (Computer Vision Center, Universitat Autònoma de Barcelona Barcelona, Spain), and Luis Herranz (Computer Vision Center, Universitat Autònoma de Barcelona Barcelona, Spain)

Structured Uncertainty Prediction Networks 5477
Garoe Dorta (University of Bath; Anthropics Technology Ltd.), Sara Vicente (Anthropics Technology Ltd.), Lourdes Agapito (University College London), Neill D. F. Campbell (University of Bath), and Ivor Simpson (Anthropics Technology Ltd.)

Between-Class Learning for Image Classification 5486
Yuji Tokozume (The University of Tokyo), Yoshitaka Ushiku (The University of Tokyo), and Tatsuya Harada (The University of Tokyo; RIKEN)

Adversarial Feature Augmentation for Unsupervised Domain Adaptation 5495
Riccardo Volpi (Pattern Analysis & Computer Vision - Istituto Italiano di Tecnologia), Pietro Morerio (Pattern Analysis & Computer Vision - Istituto Italiano di Tecnologia), Silvio Savarese (Stanford Vision and Learning Lab - Stanford University), and Vittorio Murino (Pattern Analysis & Computer Vision - Istituto Italiano di Tecnologia; Université di Verona)

Generative Image Inpainting with Contextual Attention 5505
Jiahui Yu (University of Illinois at Urbana-Champaign), Zhe Lin (Adobe Research), Jimei Yang (Adobe Research), Xiaohui Shen (Adobe Research), Xin Lu (Adobe Research), and Thomas S. Huang (University of Illinois at Urbana-Champaign)

CSGNet: Neural Shape Parser for Constructive Solid Geometry 5515
Gopal Sharma (University of Massachusetts, Amherst), Rishabh Goyal (University of Massachusetts, Amherst), Difan Liu, Evangelos Kalogerakis (University of Massachusetts, Amherst), and Subhransu Maji (University of Massachusetts, Amherst)

Conditional Image-to-Image Translation 5524
Jianxin Lin (University of Science and Technology of China), Yingce Xia (University of Science and Technology of China), Tao Qin (Microsoft Research Asia), Zhibo Chen (University of Science and Technology of China), and Tie-Yan Liu (Microsoft Research Asia)

Continuous Relaxation of MAP Inference: A Nonconvex Perspective 5533
D. Khuê Lê-Huu (CentraleSupélec, Université Paris-Saclay; Inria) and Nikos Paragios (CentraleSupélec, Université Paris-Saclay; Inria; TheraPanacea)

Feature Generating Networks for Zero-Shot Learning 5542
Yongqin Xian (Max Planck Institute for Informatics), Tobias Lorenz (Max Planck Institute for Informatics), Bernt Schiele (Max Planck Institute for Informatics), and Zeynep Akata (Max Planck Institute for Informatics; Amsterdam Machine Learning Lab Saarland Informatics Campus University of Amsterdam)

Joint Optimization Framework for Learning with Noisy Labels 5552
Daiki Tanaka (The University of Tokyo, Japan), Daiki Ikami (The University of Tokyo, Japan), Toshihiko Yamasaki (The University of Tokyo, Japan), and Kiyoharu Aizawa (The University of Tokyo, Japan)

Convolutional Image Captioning 5561
Jyoti Aneja (University of Illinois at Urbana-Champaign), Aditya Deshpande (University of Illinois at Urbana-Champaign), and Alexander G. Schwing (University of Illinois at Urbana-Champaign)

AON: Towards Arbitrarily-Oriented Text Recognition 5571
Zhanzhan Cheng (Hikvision Research Institute, China), Yangliu Xu (Tongji University, Shanghai, China), Fan Bai (Shanghai Key Lab of Intelligent Information Processing, and School of Computer Science, Fudan University, Shanghai, China), Yi Niu (Hikvision Research Institute, China), Shiliang Pu (Hikvision Research Institute, China), and Shuigeng Zhou (Shanghai Key Lab of Intelligent Information Processing, and School of Computer Science, Fudan University, Shanghai, China)

Wrapped Gaussian Process Regression on Riemannian Manifolds 5580
Anton Mallasto (University of Copenhagen) and Aasa Feragen (University of Copenhagen)

Geometry Guided Convolutional Neural Networks for Self-Supervised Video Representation Learning 5589
Chuang Gan (MIT-IBM Watson AI Lab), Boqing Gong (Tencent AI Lab), Kun Liu (BUPT), Hao Su (University of California, San Diego), and Leonidas J. Guibas (Stanford University)

DiverseNet: When One Right Answer is not Enough 5598
Michael Firman (University College London), Neill D. F. Campbell (University of Bath), Lourdes Agapito (University College London), and Gabriel J. Brostow (University College London)

Deep Face Detector Adaptation Without Negative Transfer or Catastrophic Forgetting 5608
Muhammad Abdullah Jamal (University of Central Florida Orlando, FL), Haoxiang Li (Adobe Research San Jose, CA), and Boqing Gong (Tencent AI Lab Bellevue, WA)

Analyzing Filters Toward Efficient ConvNet 5619
Takumi Kobayashi (National Institute of Advanced Industrial Science and Technology, Japan)

Regularizing Deep Networks by Modeling and Predicting Label Structure 5629
Mohammadreza Mostajabi (Toyota Technological Institute at Chicago), Michael Maire (Toyota Technological Institute at Chicago), and Gregory Shakhnarovich (Toyota Technological Institute at Chicago)

In-place Activated BatchNorm for Memory-Optimized Training of DNNs 5639
Samuel Rota Bulò (Mapillary Research), Lorenzo Porzi (Mapillary Research), and Peter Kontschieder (Mapillary Research)

DVQA: Understanding Data Visualizations via Question Answering 5648
Kushal Kafle (Rochester Institute of Technology), Brian Price (Adobe Research), Scott Cohen (Adobe Research), and Christopher Kanan (Rochester Institute of Technology)

DA-GAN: Instance-Level Image Translation by Deep Attention Generative Adversarial Networks 5657
Shuang Ma (SUNY at Buffalo), Jianlong Fu (Microsoft Research), Chang Wen Chen (SUNY at Buffalo; School of Science and Engineering, The Chinese University of Hong Kong), and Tao Mei (Microsoft Research)

Unsupervised Learning of Depth and Ego-Motion from Monocular Video Using 3D Geometric Constraints ... 5667
Reza Mahjourian (University of Texas at Austin, Google Brain), Martin Wicke (Google Brain), and Anelia Angelova (Google Brain)

FOTS: Fast Oriented Text Spotting with a Unified Network ... 5676
Xuebo Liu (SenseTime Group Ltd.), Ding Liang (SenseTime Group Ltd.), Shi Yan (SenseTime Group Ltd.), Dagui Chen (SenseTime Group Ltd.), Yu Qiao (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences), and Junjie Yan (SenseTime Group Ltd.)

Mobile Video Object Detection with Temporally-Aware Feature Maps ... 5686
Menglong Zhu (Google) and Mason Liu (Georgia Tech)

Weakly Supervised Phrase Localization with Multi-scale Anchored Transformer Network ... 5696
Fang Zhao (National University of Singapore), Jianshu Li (National University of Singapore), Jian Zhao (National University of Singapore), and Jiashi Feng (National University of Singapore)

Revisiting Oxford and Paris: Large-Scale Image Retrieval Benchmarking ... 5706
Filip Radenovic (VRG, FEE, CTU in Prague), Ahmet Iscen (VRG, FEE, CTU in Prague), Giorgos Tolias (VRG, FEE, CTU in Prague), Yannis Avrithis (Inria Rennes), and Ondrej Chum (VRG, FEE, CTU in Prague)

Cross-Dataset Adaptation for Visual Question Answering ... 5716
Fei Sha (U. of Southern California Los Angeles, CA), Hexiang Hu (U. of Southern California Los Angeles, CA), and Wei-Lun Chao (U. of Southern California Los Angeles, CA)

Globally Optimal Inlier Set Maximization for Atlanta Frame Estimation ... 5726
Kyungdon Joo (RCV Lab., KAIST), Tae-Hyun Oh (MIT CSAIL), In So Kweon (RCV Lab., KAIST), and Jean-Charles Bazin (CML, KAIST)

End-to-End Convolutional Semantic Embeddings ... 5735
Quanzeng You (Microsoft One Microsoft Way, Redmond, WA), Jiebo Luo (University of Rochester Rochester, NY), and Zhengyou Zhang (Tencent Shenzhen, China)

Referring Image Segmentation via Recurrent Refinement Networks ... 5745
Ruiyu Li (The Chinese University of Hong Kong), Kaican Li (The Chinese University of Hong Kong), Yi-Chun Kuo (The Chinese University of Hong Kong), Michelle Shu (Johns Hopkins University), Xiaojuan Qi (The Chinese University of Hong Kong), Xiaoyong Shen (YouTu Lab, Tencent), and Jiaya Jia (The Chinese University of Hong Kong; YouTu Lab, Tencent)

Two Can Play This Game: Visual Dialog with Discriminative Question Generation and Answering ... 5754
Unnat Jain (UIUC), Alexander Schwing (UIUC), and Svetlana Lazebnik (UIUC)

Generative Adversarial Learning Towards Fast Weakly Supervised Detection ... 5764
Yunhan Shen (Xiamen University, Fujian, China), Rongrong Ji (Xiamen University, Fujian, China), Shengchuan Zhang (Xiamen University, Fujian, China), Wangmeng Zuo (Harbin Institute of Technology, China), and Yan Wang (Microsoft, Redmond, USA)

A Deeper Look at Power Normalizations ... 5774
Piotr Koniusz (Data61/CSIRO; Australian National University), Hongguang Zhang (Australian National University; Data61/CSIRO), and Fatih Porikli (Australian National University)

Dimensionality's Blessing: Clustering Images by Underlying Distribution 5784
Wen-Yan Lin (Advanced Digital Sciences Center), Jian-Huang Lai (Sun Yat-Sen University), Siying Liu (Institute of Infocomm Research), and Yasuyuki Matsushita (Osaka University)

Eliminating Background-bias for Robust Person Re-identification 5794
Maoqing Tian (SenseTime Research), Shuai Yi (SenseTime Research), Hongsheng Li (Chinese University of Hong Kong), Shihua Li (Shenzhen Municipal Public Security Bureau), Xuesen Zhang (SenseTime Research), Jianping Shi (SenseTime Research), Junjie Yan (SenseTime Research), and Xiaogang Wang (Chinese University of Hong Kong)

Learning to Evaluate Image Captioning 5804
Yin Cui (Cornell University; Cornell Tech), Guandao Yang (Cornell University), Andreas Veit (Cornell University; Cornell Tech), Xun Huang (Cornell University; Cornell Tech), and Serge Belongie (Cornell University; Cornell Tech)

Single-Shot Object Detection with Enriched Semantics 5813
Zhishuai Zhang (Johns Hopkins University), Siyuan Qiao (Johns Hopkins University), Cihang Xie (Johns Hopkins University), Wei Shen (Johns Hopkins University; Shanghai University), Bo Wang (Hikvision Research), and Alan L. Yuille (Johns Hopkins University)

Low-Shot Learning with Imprinted Weights 5822
Hang Qi (UCLA), Matthew Brown (Google), and David G. Lowe (Google)

Neural Motifs: Scene Graph Parsing with Global Context 5831
Rowan Zellers (University of Washington), Mark Yatskar (University of Washington; Allen Institute for Artificial Intelligence), Sam Thomson (Carnegie Mellon University), and Yejin Choi (Paul G. Allen School of Computer Science & Engineering, University of Washington; Allen Institute for Artificial Intelligence)

Variational Autoencoders for Deforming 3D Mesh Models 5841
Qingyang Tan (Beijing Key Laboratory of Mobile Computing and Pervasive Device, Institute of Computing Technology, Chinese Academy of Sciences; School of Computer and Control Engineering, University of Chinese Academy of Sciences), Lin Gao (Beijing Key Laboratory of Mobile Computing and Pervasive Device, Institute of Computing Technology, Chinese Academy of Sciences), Yu-Kun Lai (School of Computer Science & Informatics, Cardiff University), and Shihong Xia (Beijing Key Laboratory of Mobile Computing and Pervasive Device, Institute of Computing Technology, Chinese Academy of Sciences)

Fast Monte-Carlo Localization on Aerial Vehicles Using Approximate Continuous Belief Representations..... 5851
Aditya Dhawale (Institute, Carnegie Mellon University), Kumar Shaurya Shankar (Institute, Carnegie Mellon University), and Nathan Michael (Institute, Carnegie Mellon University)

DeLS-3D: Deep Localization and Segmentation with a 3D Semantic Map 5860
Peng Wang (Baidu Research National Engineering Laboratory for Deep Learning Technology and Applications), Ruigang Yang (Baidu Research National Engineering Laboratory for Deep Learning Technology and Applications), Binbin Cao (Baidu Research National Engineering Laboratory for Deep Learning Technology and Applications), Wei Xu (Baidu Research National Engineering Laboratory for Deep Learning Technology and Applications), and Yuanqing Lin (Baidu Research National Engineering Laboratory for Deep Learning Technology and Applications)

LiDAR-Video Driving Dataset: Learning Driving Policies Effectively 5870
Yiping Chen (Fujian Key Laboratory of Sensing and Computing for Smart Cities, Xiamen University), Jingkang Wang (Shanghai Jiao Tong University), Jonathan Li (Fujian Key Laboratory of Sensing and Computing for Smart Cities, Xiamen University; University of Waterloo), Cewu Lu (Shanghai Jiao Tong University), Zhipeng Luo (Fujian Key Laboratory of Sensing and Computing for Smart Cities, Xiamen University), Han Xue (Shanghai Jiao Tong University), and Cheng Wang (Fujian Key Laboratory of Sensing and Computing for Smart Cities, Xiamen University)

Logo Synthesis and Manipulation with Clustered Generative Adversarial Networks 5879
Alexander Sage (D-ITET, ETH Zurich), Radu Timofte (D-ITET, ETH Zurich Merantix GmbH), Eirikur Agustsson (D-ITET, ETH Zurich), and Luc Van Gool (D-ITET, ETH Zurich ESAT, KU Leuven)

Egocentric Basketball Motion Planning from a Single First-Person Image 5889
Jianbo Shi (University of Pennsylvania), Aaron Chan (University of Southern California), and Gedas Bertasius (University of Pennsylvania)

Human-Centric Indoor Scene Synthesis Using Stochastic Grammar 5899
Siyuan Qi (UCLA Center for Vision, Cognition, Learning and Autonomy), Yixin Zhu (UCLA Center for Vision, Cognition, Learning and Autonomy), Siyuan Huang (UCLA Center for Vision, Cognition, Learning and Autonomy), Chenfanfu Jiang (UPenn Computer Graphics Group), and Song-Chun Zhu (UCLA Center for Vision, Cognition, Learning and Autonomy)

Rotation-Sensitive Regression for Oriented Scene Text Detection 5909
Minghui Liao (Huazhong University of Science and Technology), Zhen Zhu (Huazhong University of Science and Technology), Baoguang Shi (Huazhong University of Science and Technology), Gui-song Xia (Wuhan University), and Xiang Bai (Huazhong University of Science and Technology)

Separating Self-Expression and Visual Content in Hashtag Supervision 5919
Andreas Veit (Cornell University), Maximilian Nickel (Facebook AI Research), Serge Belongie (Cornell University), and Laurens van der Maaten (Facebook AI Research)

Distort-and-Recover: Color Enhancement Using Deep Reinforcement Learning 5928
Jongchan Park (Lunit Inc.), Joon-Young Lee (Adobe Research), Donggeun Yoo (Lunit Inc.), Donggeun Yoo (Korea Advanced Institute of Science and Technology (KAIST)), and In So Kweon (Korea Advanced Institute of Science and Technology (KAIST))

Oral and Spotlight 2-2A

O2-2A: Video Analytics

Im2Flow: Motion Hallucination from Static Images for Action Recognition 5937
Ruohan Gao (UT Austin), Bo Xiong (UT Austin), and Kristen Grauman (UT Austin)

Finding "It": Weakly-Supervised Reference-Aware Visual Grounding in Instructional Videos 5948
De-An Huang (Stanford University), Shyamal Buch (Stanford University), Lucio Dery (Stanford University), Animesh Garg (Stanford University), Li Fei-Fei (Stanford University), and Juan Carlos Niebles (Stanford University)

Actor and Action Video Segmentation from a Sentence 5958
Kirill Gavrilyuk (QUVA Lab, University of Amsterdam), Amir Ghodrati (QUVA Lab, University of Amsterdam), Zhenyang Li (QUVA Lab, University of Amsterdam), and Cees G. M. Snoek (QUVA Lab, University of Amsterdam)

Egocentric Activity Recognition on a Budget 5967
Rafael Possas (University of Sydney), Sheila Pinto Caceres (University of Sydney), and Fabio Ramos (University of Sydney)

S2-2A

CNN in MRF: Video Object Segmentation via Inference in a CNN-Based Higher-Order Spatio-Temporal MRF 5977
Linchao Bao (Tencent AI Lab), Baoyuan Wu (Tencent AI Lab), and Wei Liu (Tencent AI Lab)

Action Sets: Weakly Supervised Action Segmentation Without Ordering Constraints 5987
Alexander Richard (University of Bonn, Germany), Hilde Kuehne (University of Bonn, Germany), and Juergen Gall (University of Bonn, Germany)

Low-Latency Video Semantic Segmentation 5997
Yule Li (Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China), Jianping Shi (SenseTime Research), and Dahua Lin (The Chinese University of Hong Kong)

Fine-Grained Video Captioning for Sports Narrative 6006
Huanyu Yu (Shanghai Jiao Tong University), Shuo Cheng (Shanghai Jiao Tong University), Bingbing Ni (Shanghai Jiao Tong University), Minsi Wang (Shanghai Jiao Tong University), Jian Zhang (Shanghai Jiao Tong University), and Xiaokang Yang (Shanghai Jiao Tong University)

End-to-End Learning of Motion Representation for Video Understanding 6016
Lijie Fan (Tsinghua University, Beijing, China), Wenbing Huang (Tencent AI Lab), Chuang Gan (MIT-Watson Lab), Stefano Ermon (Stanford University), Boqing Gong (Tencent AI Lab), and Junzhou Huang (Tencent AI Lab)

Compressed Video Action Recognition 6026
Chao-Yuan Wu (The University of Texas at Austin; Amazon), Manzil Zaheer (Carnegie Mellon University; Amazon), Hexiang Hu (University of Southern California; Amazon), R. Manmatha (A9), Alexander J. Smola (Amazon), and Philipp Krähenbühl (The University of Texas at Austin)

Features for Multi-target Multi-camera Tracking and Re-identification 6036
Ergys Ristani (Duke University Durham, NC, USA) and Carlo Tomasi (Duke University Durham, NC, USA)

AVA: A Video Dataset of Spatio-Temporally Localized Atomic Visual Actions 6047
Chunhui Gu (Google Research), Chen Sun (Google Research), David A. Ross (Google Research), Carl Vondrick (Google Research), Caroline Pantofaru (Google Research), Yeqing Li (Google Research), Sudheendra Vijayanarasimhan (Google Research), George Toderici (Google Research), Susanna Ricco (Google Research), Rahul Sukthankar (Google Research), Cordelia Schmid (Inria, Laboratoire Jean Kuntzmann, Grenoble, France; Google Research), and Jitendra Malik (University of California at Berkeley, USA; Google Research)

Who's Better? Who's Best? Pairwise Deep Ranking for Skill Determination 6057
Hazel Doughty (University of Bristol, Bristol), Dima Damen (University of Bristol, Bristol), and Walterio Mayol-Cuevas (University of Bristol, Bristol)

MX-LSTM: Mixing Tracklets and Vislets to Jointly Forecast Trajectories and Head Poses 6067
Irtiza Hasan (University of Verona (UNIVR); OSRAM GmbH), Francesco Setti (University of Verona (UNIVR)), Theodore Tsesmelis (University of Verona (UNIVR); OSRAM GmbH; Istituto Italiano di Tecnologia (IIT)), Alessio Del Bue (Istituto Italiano di Tecnologia (IIT)), Fabio Galasso (OSRAM GmbH), and Marco Cristani (University of Verona (UNIVR))

Oral and Spotlight 2-2B

O2-2B: Object Recognition & Scene Understanding III

Bottom-Up and Top-Down Attention for Image Captioning and Visual Question Answering 6077
Peter Anderson (Australian National University), Xiaodong He (JD AI Research), Chris Buehler (Microsoft Research), Damien Teney (University of Adelaide), Mark Johnson (Macquarie University), Stephen Gould (Australian National University), and Lei Zhang (Microsoft Research)

Improved Fusion of Visual and Language Representations by Dense Symmetric Co-attention for Visual Question Answering 6087
Duy-Kien Nguyen (Tohoku University) and Takayuki Okatani (Tohoku University; RIKEN Center for AIP)

FLIPDIAL: A Generative Model for Two-Way Visual Dialogue 6097
Puneet K. Dokania (University of Oxford), Philip H.S. Torr (University of Oxford), N. Siddharth (University of Oxford), and Daniela Massiceti (University of Oxford)

Are You Talking to Me? Reasoned Visual Dialog Generation Through Adversarial Learning 6106
Qi Wu (Australian Centre for Robotic Vision, The University of Adelaide, Australia), Peng Wang (School of Computer Science and Engineering, Northwestern Polytechnical University), Chunhua Shen (Australian Centre for Robotic Vision, The University of Adelaide, Australia), Ian Reid (Australian Centre for Robotic Vision, The University of Adelaide, Australia), and Anton van den Hengel (Australian Centre for Robotic Vision, The University of Adelaide, Australia)

S2-2B

Visual Question Generation as Dual Task of Visual Question Answering 6116
Yikang Li (The Chinese University of Hong Kong, Hong Kong, China), Nan Duan (Microsoft Research Asia, China), Bolei Zhou (Massachusetts Institute of Technology, USA), Xiao Chu (The Chinese University of Hong Kong, Hong Kong, China), Wanli Ouyang (University of Sydney, Australia), Xiaogang Wang (The Chinese University of Hong Kong, Hong Kong, China), and Ming Zhou (Microsoft Research Asia, China)

Unsupervised Textual Grounding: Linking Words to Image Concepts 6125
Raymond A. Yeh (University of Illinois at Urbana-Champaign), Minh N. Do (University of Illinois at Urbana-Champaign), and Alexander G. Schwing (University of Illinois at Urbana-Champaign)

Focal Visual-Text Attention for Visual Question Answering 6135
Junwei Liang (Carnegie Mellon University), Lu Jiang (Google Inc.), Liangliang Cao (HelloVera AI), Li-Jia Li (Google Inc.), and Alexander Hauptmann (Carnegie Mellon University)

SeGAN: Segmenting and Generating the Invisible 6144
Kiana Ehsani (University of Washington), Roozbeh Mottaghi (Allen Institute for AI (AI2)), and Ali Farhadi (University of Washington; Allen Institute for AI (AI2))

Cascade R-CNN: Delving Into High Quality Object Detection 6154
Zhaowei Cai (UC San Diego) and Nuno Vasconcelos (UC San Diego)

Learning Semantic Concepts and Order for Image and Sentence Matching 6163
Yan Huang (Center for Research on Intelligent Perception and Computing (CRIPAC), National Laboratory of Pattern Recognition (NLPR); University of Chinese Academy of Sciences (UCAS)), Qi Wu (School of Computer Science, The University of Adelaide), Chunfeng Song (Center for Research on Intelligent Perception and Computing (CRIPAC), National Laboratory of Pattern Recognition (NLPR); University of Chinese Academy of Sciences (UCAS)), and Liang Wang (Center for Research on Intelligent Perception and Computing (CRIPAC), National Laboratory of Pattern Recognition (NLPR); Center for Excellence in Brain Science and Intelligence Technology (CEBSIT), Institute of Automation, Chinese Academy of Sciences (CASIA); University of Chinese Academy of Sciences (UCAS))

Functional Map of the World 6172
Gordon Christie (The Johns Hopkins University Applied Physics Laboratory), Neil Fendley (The Johns Hopkins University Applied Physics Laboratory), James Wilson (DigitalGlobe), and Ryan Mukherjee (The Johns Hopkins University Applied Physics Laboratory)

MegDet: A Large Mini-Batch Object Detector 6181
Chao Peng, Tete Xiao (Peking University), Zeming Li (Tsinghua University, China), Yuning Jiang, Xiangyu Zhang, Kai Jia, Gang Yu, and Jian Sun

Learning Globally Optimized Object Detector via Policy Gradient 6190
Yongming Rao (Tsinghua University, China; State Key Lab of Intelligent Technologies and Systems, Tsinghua University, China; Beijing National Research Center for Information Science and Technology), Dahua Lin (The Chinese University of Hong Kong), Jiwen Lu (Tsinghua University, China; State Key Lab of Intelligent Technologies and Systems, Tsinghua University, China; Beijing National Research Center for Information Science and Technology), and Jie Zhou (Tsinghua University, China; State Key Lab of Intelligent Technologies and Systems, Tsinghua University, China; Beijing National Research Center for Information Science and Technology)

Photographic Text-to-Image Synthesis with a Hierarchically-Nested Adversarial Network 6199
Zizhao Zhang (University of Florida), Yuanpu Xie (University of Florida), and Lin Yang (University of Florida)

Oral and Spotlight 2-2C

O2-2C: Computational Photography

Illuminant Spectra-Based Source Separation Using Flash Photography 6209
Zhuo Hui (Carnegie Mellon University), Kalyan Sunkavalli (Adobe Research), Sunil Hadap (Adobe Research), and Aswin C. Sankaranarayanan (Carnegie Mellon University)

Trapping Light for Time of Flight 6219
Ruilin Xu (Columbia University), Shree K. Nayar (Columbia University), and Mohit Gupta (University of Wisconsin-Madison)

The Perception-Distortion Tradeoff 6228
Yochai Blau (Technion-Israel Institute of Technology, Haifa, Israel) and Tomer Michaeli (Technion-Israel Institute of Technology, Haifa, Israel)

S2-2C

Label Denoising Adversarial Network (LDAN) for Inverse Lighting of Faces 6238
Hao Zhou (University of Maryland, College Park, MD, USA), Jin Sun (University of Maryland, College Park, MD, USA), Yaser Yacoob (University of Maryland, College Park, MD, USA), and David W. Jacobs (University of Maryland, College Park, MD, USA)

Optimal Structured Light a la Carte 6248
Parsa Mirdehghan (University of Toronto), Wenzheng Chen (University of Toronto), and Kiriakos N. Kutulakos (University of Toronto)

Tracking Multiple Objects Outside the Line of Sight Using Speckle Imaging 6258
Brandon M. Smith (University of Wisconsin-Madison), Matthew O'Toole (Stanford University), and Mohit Gupta (University of Wisconsin-Madison)

Inferring Light Fields from Shadows 6267
Manel Baradad (Massachusetts Institute of Technology), Vickie Ye (Massachusetts Institute of Technology), Adam B. Yedidia (Massachusetts Institute of Technology), Frédo Durand (Massachusetts Institute of Technology), William T. Freeman (Massachusetts Institute of Technology; Google Research), Gregory W. Wornell (Massachusetts Institute of Technology), and Antonio Torralba (Massachusetts Institute of Technology)

Modifying Non-local Variations Across Multiple Views 6276
Tal Tlusty (Technion, Israel), Tomer Michaeli (Technion, Israel), Tali Dekel (Google Research), and Lihi Zelnik-Manor (Technion, Israel)

Robust Video Content Alignment and Compensation for Rain Removal in a CNN Framework 6286
Jie Chen (Nanyang Technological University, Singapore), Cheen-Hau Tan (Nanyang Technological University, Singapore), Junhui Hou (City University of Hong Kong), Lap-Pui Chau (Nanyang Technological University, Singapore), and He Li (Singapore Technologies Dynamics Pte Ltd)

SfSNet: Learning Shape, Reflectance and Illuminance of Faces 'in the Wild' 6296
Soumyadip Sengupta (University of Maryland, College Park), Angjoo Kanazawa (University of California, Berkeley), Carlos D. Castillo (University of Maryland, College Park), and David W. Jacobs (University of Maryland, College Park)

Deep Photo Enhancer: Unpaired Learning for Image Enhancement from Photographs with GANs 6306
Yu-Sheng Chen (National Taiwan University), Yu-Ching Wang (National Taiwan University), Man-Hsin Kao (National Taiwan University), and Yung-Yu Chuang (National Taiwan University)

LIME: Live Intrinsic Material Estimation 6315
Abhimitra Meka (MPI Informatics; Saarland Informatics Campus), Maxim Maximov (MPI Informatics; Saarland Informatics Campus), Michael Zollhöfer (MPI Informatics; Saarland Informatics Campus; Stanford University), Avishek Chatterjee (MPI Informatics; Saarland Informatics Campus), Hans-Peter Seidel (MPI Informatics; Saarland Informatics Campus), Christian Richardt (University of Bath), and Christian Theobalt (MPI Informatics; Saarland Informatics Campus)

Learning to Detect Features in Texture Images 6325
Linguang Zhang (Princeton University) and Szymon Rusinkiewicz (Princeton University)

Learning to Extract a Video Sequence from a Single Motion-Blurred Image 6334
Meiguang Jin (University of Bern, Switzerland), Givi Meishvili (University of Bern, Switzerland), and Paolo Favaro (University of Bern, Switzerland)

Lose The Views: Limited Angle CT Reconstruction Via Implicit Sinogram Completion 6343
Rushil Anirudh (Lawrence Livermore National Laboratory.), Hyojin Kim (Lawrence Livermore National Laboratory.), Jayaraman J. Thiagarajan (Lawrence Livermore National Laboratory.), K. Aditya Mohan (Lawrence Livermore National Laboratory.), Kyle Champley (Lawrence Livermore National Laboratory.), and Timo Bremer (Lawrence Livermore National Laboratory.)

A Common Framework for Interactive Texture Transfer 6353
Yifang Men (Institute of Computer Science and Technology, Peking University, China), Zhouhui Lian (Institute of Computer Science and Technology, Peking University, China), Yingmin Tang (Institute of Computer Science and Technology, Peking University, China), and Jianguo Xiao (Institute of Computer Science and Technology, Peking University, China)

Poster Session Poster 2-3

AMNet: Memorability Estimation with Attention 6363
Jiri Fajtl (Kingston University, London), Vasileios Argyriou (Kingston University, London), Dorothy Monekosso (Leeds Beckett University, Leeds), and Paolo Remagnino (Kingston University, London)

Blind Predicting Similar Quality Map for Image Quality Assessment 6373
Da Pan (Communication University of China), Ping Shi (Communication University of China), Ming Hou (Communication University of China), Zefeng Ying (Communication University of China), Sizhe Fu (Communication University of China), and Yuan Zhang (Communication University of China)

Deep End-to-End Time-of-Flight Imaging 6383
Shuochen Su (UBC), Felix Heide (Stanford University), Gordon Wetzstein (Stanford University), and Wolfgang Heidrich (KAUST)

Aperture Supervision for Monocular Depth Estimation 6393
Pratul P. Srinivasan (UC Berkeley), Rahul Garg (Google Research), Neal Wadhwa (Google Research), Ren Ng (UC Berkeley), and Jonathan T. Barron (Google Research)

Seeing Temporal Modulation of Lights from Standard Cameras 6402
Naoki Sakakibara (Nagoya Institute of Technology), Fumihiko Sakaue (Nagoya Institute of Technology), and Jun Sato (Nagoya Institute of Technology)

Statistical Tomography of Microscopic Life 6411
Aviad Levis (Technion - Israel Institute of Technology, Haifa, Israel), Yoav Y. Schechner (Technion - Israel Institute of Technology, Haifa, Israel), and Ronen Talmon (Technion - Israel Institute of Technology, Haifa, Israel)

Divide and Conquer for Full-Resolution Light Field Deblurring 6421
M. R. Mahesh Mohan (Indian Institute of Technology Madras) and A. N. Rajagopalan (Indian Institute of Technology Madras)

Multispectral Image Intrinsic Decomposition via Subspace Constraint 6430
Qian Huang (Nanjing University Nanjing, China), Weixin Zhu (Nanjing University Nanjing, China), Yang Zhao (Nanjing University Nanjing, China), Linsen Chen (Nanjing University Nanjing, China), Yao Wang (New York University, USA), Tao Yue (Nanjing University Nanjing, China), and Xun Cao (Nanjing University, Nanjing, China)

Improving Color Reproduction Accuracy on Cameras 6440
Hakki Can Karaimer (York University, Toronto) and Michael S. Brown (York University, Toronto)

A Closer Look at Spatiotemporal Convolutions for Action Recognition 6450
Du Tran (Facebook Research), Heng Wang (Facebook Research), Lorenzo Torresani (Facebook Research; Dartmouth College), Jamie Ray (Facebook Research), Yann LeCun (Facebook Research), and Manohar Paluri (Facebook Research)

Inferring Shared Attention in Social Scene Videos 6460
Lifeng Fan (Center for Vision, Cognition, Learning and Autonomy, UCLA, Los Angeles, USA), Yixin Chen (Center for Vision, Cognition, Learning and Autonomy, UCLA, Los Angeles, USA), Ping Wei (Xi'an Jiaotong University, Xi' an, China; Center for Vision, Cognition, Learning and Autonomy, UCLA, Los Angeles, USA), Wenguan Wang (Beijing Institute of Technology, Beijing, China; Center for Vision, Cognition, Learning and Autonomy, UCLA, Los Angeles, USA), and Song-Chun Zhu (Center for Vision, Cognition, Learning and Autonomy, UCLA, Los Angeles, USA)

Making Convolutional Networks Recurrent for Visual Sequence Learning 6469
Xiaodong Yang (NVIDIA), Pavlo Molchanov (NVIDIA), and Jan Kautz (NVIDIA)

Real-World Anomaly Detection in Surveillance Videos 6479
Waqas Sultani (Information Technology University, Pakistan), Chen Chen (University of Central Florida, Orlando, FL), and Mubarak Shah (University of Central Florida, Orlando, FL)

Viewpoint-Aware Attentive Multi-view Inference for Vehicle Re-identification 6489
Yi Zhouy (University of East Anglia) and Ling Shao (University of East Anglia)

Efficient Video Object Segmentation via Network Modulation 6499
Linjie Yang (Snap Inc.), Yanran Wang (Northwestern University), Xuehan Xiong (Google Inc.), Jianchao Yang (Snap Inc.), and Aggelos K. Katsaggelos (Northwestern University)

Weakly-Supervised Action Segmentation with Iterative Soft Boundary Assignment 6508
Chenliang Xu (University of Rochester) and Li Ding (Massachusetts Institute of Technology)

Depth-Aware Stereo Video Retargeting 6517
Bing Li (University of Southern California, Los Angeles, California, USA), Chia-Wen Lin (National Tsing Hua University), Boxin Shi (Peking University, Beijing, China), Tiejun Huang (Peking University, Beijing, China), Wen Gao (Peking University, Beijing, China), and C.-C. Jay Kuo (University of Southern California, Los Angeles, California, USA)

Instance Embedding Transfer to Unsupervised Video Object Segmentation 6526
Siyang Li (University of Southern California; Google AI Perception), Bryan Seybold (Google AI Perception), Alexey Vorobyov (Google AI Perception), Alireza Fathi (Google AI Perception), Qin Huang (University of Southern California), and C.-C. Jay Kuo (University of Southern California)

Future Frame Prediction for Anomaly Detection - A New Baseline 6536
Wen Liu (ShanghaiTech University), Weixin Luo (ShanghaiTech University), Dongze Lian (ShanghaiTech University), and Shenghua Gao (ShanghaiTech University)

Can Spatiotemporal 3D CNNs Retrace the History of 2D CNNs and ImageNet? 6546
Kensho Hara (National Institute of Advanced Industrial Science and Technology (AIST)), Hirokatsu Kataoka (National Institute of Advanced Industrial Science and Technology (AIST)), and Yutaka Satoh (National Institute of Advanced Industrial Science and Technology (AIST))

Dynamic Video Segmentation Network 6556
Yu-Syuan Xu (National Tsing Hua Uiversity), Tsu-Jui Fu (National Tsing Hua Uiversity), Hsuan-Kung Yang (National Tsing Hua Uiversity), and Chun-Yi Lee (National Tsing Hua Uiversity)

Recognize Actions by Disentangling Components of Dynamics 6566
Yue Zhao (CUHK - SenseTime Joint Lab, The Chinese University of Hong Kong), Yuanjun Xiong (CUHK - SenseTime Joint Lab, The Chinese University of Hong Kong; Amazon Rekognition), and Dahua Lin (CUHK - SenseTime Joint Lab, The Chinese University of Hong Kong)

Motion-Appearance Co-memory Networks for Video Question Answering 6576
Jiyang Gao (University of Southern California), Runzhou Ge (University of Southern California), Kan Chen (University of Southern California), and Ram Nevatia (University of Southern California)

Learning to Understand Image Blur 6586
Shanghang Zhang (Carnegie Mellon University; ISR - IST, Universidade de Lisboa), Xiaohui Shen (Adobe Research), Zhe Lin (Adobe Research), Radomí-r Mech (Adobe Research), João P. Costeira (ISR - IST, Universidade de Lisboa), and Jose M.F. Moura

Dense Decoder Shortcut Connections for Single-Pass Semantic Segmentation 6596
Piotr Bilinski (University of Oxford) and Victor Prisacariu (University of Oxford)

Generative Adversarial Image Synthesis with Decision Tree Latent Controller 6606
Takuhiro Kaneko (NTT Communication Science Laboratories, NTT Corporation), Kaoru Hiramatsu (NTT Communication Science Laboratories, NTT Corporation), and Kunio Kashino (NTT Communication Science Laboratories, NTT Corporation)

Learning a Discriminative Prior for Blind Image Deblurring 6616
Lerenhan Li (Huazhong University of Science and Technology; University of California, Merced), Jinshan Pan (Nanjing University of Science and Technology), Wei-Sheng Lai (University of California, Merced), Changxin Gao (Huazhong University of Science and Technology), Nong Sang (Huazhong University of Science and Technology), and Ming-Hsuan Yang (University of California, Merced)

Frame-Recurrent Video Super-Resolution 6626
Mehdi S. M. Sajjadi (Max Planck Institute for Intelligent Systems; Google), Raviteja Vemulapalli (Google), and Matthew Brown (Google)

Discovering Point Lights with Intensity Distance Fields 6635
Michael F. Cohen (Facebook, Inc. Seattle, WA), Edward Zhang (University of Washington Seattle, WA), and Brian Curless (University of Washington Seattle, WA)

Video Rain Streak Removal by Multiscale Convolutional Sparse Coding 6644
Minghan Li (Xian Jiaotong University), Qi Xie (Xian Jiaotong University), Qian Zhao (Xian Jiaotong University), Wei Wei (Xian Jiaotong University), Shuhang Gu (Computer Vision Lab, ETH Zurich), Jing Tao (Xian Jiaotong University), and Deyu Meng (Xian Jiaotong University)

Stereoscopic Neural Style Transfer 6654
Dongdong Chen (University of Science and Technology of China), Lu Yuan (Microsoft Research), Jing Liao (Microsoft Research), Nenghai Yu (University of Science and Technology of China), and Gang Hua (Microsoft Research)

Multi-Frame Quality Enhancement for Compressed Video 6664
Ren Yang (Beihang University, Beijing, China), Mai Xu (Beihang University, Beijing, China), Zulin Wang (Beihang University, Beijing, China), and Tianyi Li (Beihang University, Beijing, China)

CNN Based Learning Using Reflection and Retinex Models for Intrinsic Image Decomposition 6674
Anil S. Baslamisli (Informatics Institute University of Amsterdam), Hoang-An Le (Informatics Institute University of Amsterdam), and Theo Gevers (Informatics Institute University of Amsterdam)

Image Restoration by Estimating Frequency Distribution of Local Patches 6684
Nojun Kwak (Seoul National University), Jaeyoung Yoo (Seoul National University; NALBI Inc.), and Sang-ho Lee (Seoul National University)

Latent RANSAC 6693
Simon Korman (Weizmann Institute of Science, Israel General Motors, Israel) and Roee Litman (Weizmann Institute of Science, Israel General Motors, Israel)

Two-Stream Convolutional Networks for Dynamic Texture Synthesis 6703
Matthew Tesfaldet (York University), Marcus A. Brubaker (York University), and Konstantinos G. Derpanis (Ryerson University)

Towards Open-Set Identity Preserving Face Synthesis 6713
Jianmin Bao (University of Science and Technology of China), Dong Chen (Microsoft Research), Fang Wen (Microsoft Research), Houqiang Li (University of Science and Technology of China), and Gang Hua (Microsoft Research)

A Revised Underwater Image Formation Model 6723
Derya Akkaynak (University of Haifa) and Tali Treibitz (University of Haifa)

Graph-Cut RANSAC 6733
Daniel Barath (Machine Perception Research Laboratory, MTA SZTAKI, Budapest, Hungary; Centre for Machine Perception, Czech Technical University, Prague, Czech Republic) and Jiri Matas (Centre for Machine Perception, Czech Technical University, Prague, Czech Republic)

Temporal Deformable Residual Networks for Action Segmentation in Videos 6742
Peng Lei (Oregon State University Corvallis, OR, USA) and Sinisa Todorovic (Oregon State University Corvallis, OR, USA)

Weakly Supervised Action Localization by Sparse Temporal Pooling Network 6752
Phuc Nguyen (University of California Irvine, CA, USA), Bohyung Han (Seoul National University Seoul, Korea), Ting Liu (Google Venice, CA, USA), and Gautam Prasad (Google Venice, CA, USA)

PoseFlow: A Deep Motion Representation for Understanding Human Behaviors in Videos 6762
Dingwen Zhang (Northwestern Polytechincal University; Carnegie Mellon University), Guangyu Guo (Northwestern Polytechincal University), Dong Huang (Carnegie Mellon University), and Junwei Han (Northwestern Polytechincal University)

FFNet: Video Fast-Forwarding via Reinforcement Learning 6771
Shuyue Lan (Northwestern University, USA), Rameswar Panda (University of California, Riverside), Qi Zhu (Northwestern University, USA), and Amit K. Roy-Chowdhury (University of California, Riverside)

Multi-shot Pedestrian Re-identification via Sequential Decision Making 6781
Jianfu Zhang (Shanghai Jiao Tong University), Naiyan Wang (TuSimple), and Liqing Zhang (Shanghai Jiao Tong University)

Attend and Interact: Higher-Order Object Interactions for Video Understanding 6790
Chih-Yao Ma (Georgia Institute of Technology), Asim Kadav (NEC Laboratories America), Iain Melvin (NEC Laboratories America), Zsolt Kira (Georgia Tech Research Institute), Ghassan AlRegib (Georgia Institute of Technology), and Hans Peter Graf (NEC Laboratories America)

Where and Why are They Looking? Jointly Inferring Human Attention and Intentions in Complex Tasks 6801
Ping Wei (Xi'an Jiaotong University, China; University of California, Los Angeles), Yang Liu (University of California, Los Angeles), Tianmin Shu (University of California, Los Angeles), Nanning Zheng (Xi'an Jiaotong University, China), and Song-Chun Zhu (University of California, Los Angeles)

Fully Convolutional Adaptation Networks for Semantic Segmentation 6810
Yiheng Zhang (University of Science and Technology of China, Hefei, China), Zhaofan Qiu (University of Science and Technology of China, Hefei, China), Ting Yao (Microsoft Research, Beijing, China), Dong Liu (University of Science and Technology of China, Hefei, China), and Tao Mei (Microsoft Research, Beijing, China)

Semantic Video Segmentation by Gated Recurrent Flow Propagation 6819
David Nilsson (Lund University) and Cristian Sminchisescu (Lund University; Institute of Mathematics of the Romanian Academy)

Interpretable Video Captioning via Trajectory Structured Localization 6829
Xian Wu (Sun Yat-sen University), Guanbin Li (Sun Yat-sen University), Qingxing Cao (Sun Yat-sen University), Qingge Ji (Sun Yat-sen University), and Liang Lin (Sun Yat-sen University; SenseTime Group Limited)

Deep Hashing via Discrepancy Minimization 6838
Zhixiang Chen (Tsinghua University, China; State Key Lab of Intelligent Technologies and Systems, Tsinghua University, China; Beijing National Research Center for Information Science and Technology, China), Xin Yuan (Tsinghua University, China; State Key Lab of Intelligent Technologies and Systems, Tsinghua University, China; Beijing National Research Center for Information Science and Technology, China), Jiwen Lu (Tsinghua University, China; State Key Lab of Intelligent Technologies and Systems, Tsinghua University, China; Beijing National Research Center for Information Science and Technology, China), Qi Tian (University of Texas at San Antonio, USA), and Jie Zhou (Tsinghua University, China; State Key Lab of Intelligent Technologies and Systems, Tsinghua University, China; Beijing National Research Center for Information Science and Technology, China)

ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices 6848
Xiangyu Zhang (Megvii Inc (Face++)), Xinyu Zhou (Megvii Inc (Face++)), Mengxiao Lin (Megvii Inc (Face++)), and Jian Sun (Megvii Inc (Face++))

Zero-Shot Recognition via Semantic Embeddings and Knowledge Graphs 6857
Xiaolong Wang (The Robotics Institute, Carnegie Mellon University), Yufei Ye (The Robotics Institute, Carnegie Mellon University), and Abhinav Gupta (The Robotics Institute, Carnegie Mellon University)

Referring Relationships 6867
Ranjay Krishna (Stanford University), Ines Chami (Stanford University), Michael Bernstein (Stanford University), and Li Fei-Fei (Stanford University)

Improving Object Localization with Fitness NMS and Bounded IoU Loss 6877
Lachlan Tychsen-Smith (CSIRO (Data61)) and Lars Petersson (CSIRO (Data61))

End-to-End Deep Kronecker-Product Matching for Person Re-identification 6886
Yantao Shen (CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong), Tong Xiao (CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong), Hongsheng Li (CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong), Shuai Yi (SenseTime Research), and Xiaogang Wang (CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong)

Semantic Visual Localization 6896
Johannes L. Schönberger (ETH Zurich), Marc Pollefeys (ETH Zurich; Microsoft), Andreas Geiger (ETH Zurich; Autonomous Vision Group, MPI Tübingen), and Torsten Sattler (ETH Zurich)

Objects as Context for Detecting Their Semantic Parts 6907
Abel Gonzalez-Garcia (University of Edinburgh), Davide Modolo (University of Edinburgh), and Vittorio Ferrari (University of Edinburgh)

End-to-End Weakly-Supervised Semantic Alignment 6917
Ignacio Rocco (DI ENS; Inria), Relja Arandjelovic (DeepMind), and Josef Sivic (DI ENS; Inria; CIIRC, CTU in Prague)

Dynamic Zoom-in Network for Fast Object Detection in Large Images 6926
Mingfei Gao (University of Maryland, College Park), Ruichi Yu (University of Maryland, College Park), Ang Li (DeepMind), Vlad I. Morariu (Adobe Research), and Larry S. Davis (University of Maryland, College Park)

Learning Markov Clustering Networks for Scene Text Detection 6936
Zichuan Liu (Nanyang Technological University, Singapore), Guosheng Lin (Nanyang Technological University, Singapore), Sheng Yang (Nanyang Technological University, Singapore), Jiashi Feng (National University of Singapore, Singapore), Weisi Lin (Nanyang Technological University, Singapore), and Wang Ling Goh (Nanyang Technological University, Singapore)

Deep Reinforcement Learning of Region Proposal Networks for Object Detection 6945
Aleksis Pirinen (Lund University) and Cristian Sminchisescu (Lund University; Institute of Mathematics of the Romanian Academy)

Beyond Holistic Object Recognition: Enriching Image Understanding with Part States 6955
Cewu Lu (Shanghai Jiao Tong University), Hao Su (University of California San Diego), Yonglu Li (Shanghai Jiao Tong University), Yongyi Lu (Hong Kong University of Science and Technology), Li Yi (Stanford University), Chi-Keung Tang (Hong Kong University of Science and Technology), and Leonidas J. Guibas (Stanford University)

Discriminability Objective for Training Descriptive Captions 6964
Ruotian Luo (TTI-Chicago), Gregory Shakhnarovich (TTI-Chicago), Scott Cohen (Adobe Research), and Brian Price (Adobe Research)

Visual Question Answering with Memory-Augmented Networks 6975
Chao Ma (Australian Institute for Machine Learning, and Australian Centre for Robotic Vision, The University of Adelaide), Chunhua Shen (Australian Institute for Machine Learning, and Australian Centre for Robotic Vision, The University of Adelaide), Anthony Dick (Australian Institute for Machine Learning, and Australian Centre for Robotic Vision, The University of Adelaide), Qi Wu (Australian Institute for Machine Learning, and Australian Centre for Robotic Vision, The University of Adelaide), Peng Wang (Australian Institute for Machine Learning, and Australian Centre for Robotic Vision, The University of Adelaide), Anton van den Hengel (Australian Institute for Machine Learning, and Australian Centre for Robotic Vision, The University of Adelaide), and Ian Reid (Australian Institute for Machine Learning, and Australian Centre for Robotic Vision, The University of Adelaide)

Structure Inference Net: Object Detection Using Scene-Level Context and Instance-Level Relationships..... 6985
Yong Liu (Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China; University of Chinese Academy of Sciences, Beijing, China), Ruiping Wang (Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China; University of Chinese Academy of Sciences, Beijing, China; Cooperative Medianet Innovation Center, China), Shiguang Shan (Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China; University of Chinese Academy of Sciences, Beijing, China; Cooperative Medianet Innovation Center, China), and Xilin Chen (Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China; University of Chinese Academy of Sciences, Beijing, China; Cooperative Medianet Innovation Center, China)

Occluded Pedestrian Detection Through Guided Attention in CNNs 6995
Shanshan Zhang (Key Lab of Intelligent Perception and Systems for High-Dimensional Information of Ministry of Education), Jian Yang (Nanjing University of Science and Technology, China), and Bernt Schiele (Max Planck Institute for Informatics, Saarland Informatics Campus, Germany)

Reward Learning from Narrated Demonstrations 7004
Hsiao-Yu Tung (Carnegie Mellon University, Pittsburgh, PA), Adam W. Harley (Carnegie Mellon University, Pittsburgh, PA), Liang-Kang Huang (Carnegie Mellon University, Pittsburgh, PA), and Katerina Fragkiadaki (Carnegie Mellon University, Pittsburgh, PA)

Weakly-Supervised Semantic Segmentation Network with Deep Seeded Region Growing 7014
Zilong Huang (Huazhong University of Science and Technology), Xinggang Wang (Huazhong University of Science and Technology), Jiasi Wang (Huazhong University of Science and Technology), Wenyu Liu (Huazhong University of Science and Technology), and Jingdong Wang (Microsoft Research Asia)

PoTion: Pose MoTion Representation for Action Recognition 7024
Vasileios Choutas (Inria; NAVER LABS Europe), Philippe Weinzaepfel (NAVER LABS Europe), Jérôme Revaud (NAVER LABS Europe), and Cordelia Schmid (Inria)

Bilateral Ordinal Relevance Multi-instance Regression for Facial Action Unit Intensity Estimation 7034
Yong Zhang (National Laboratory of Pattern Recognition, CASIA; University of Chinese Academy of Sciences), Rui Zhao (Rensselaer Polytechnic Institute), Weiming Dong (National Laboratory of Pattern Recognition, CASIA), Bao-Gang Hu (National Laboratory of Pattern Recognition, CASIA), and Qiang Ji (Rensselaer Polytechnic Institute)

Pulling Actions out of Context: Explicit Separation for Effective Combination 7044
Yang Wang (Stony Brook University, Stony Brook, NY, USA) and Minh Hoai (Stony Brook University, Stony Brook, NY, USA)

Dynamic Feature Learning for Partial Face Recognition 7054
Lingxiao He (CRIPAC & NLPR, CASIA; University of Chinese Academy of Sciences, Beijing, P.R. China), Haiqing Li (CRIPAC & NLPR, CASIA; University of Chinese Academy of Sciences, Beijing, P.R. China), Qi Zhang (CRIPAC & NLPR, CASIA; University of Chinese Academy of Sciences, Beijing, P.R. China), and Zhenan Sun (CRIPAC & NLPR, CASIA; University of Chinese Academy of Sciences, Beijing, P.R. China; Center for Excellence in Brain Science and Intelligence Technology, CAS)

Exploiting Transitivity for Learning Person Re-identification Models on a Budget 7064
Sourya Roy (University of California, Riverside, CA, USA), Sujoy Paul (University of California, Riverside, CA, USA), Neal E. Young (University of California, Riverside, CA, USA), and Amit K. Roy-Chowdhury (University of California, Riverside, CA, USA)

Deep Spatial Feature Reconstruction for Partial Person Re-identification: Alignment-free Approach 7073
Lingxiao He (CRIPAC & NLPR, CASIA; University of Chinese Academy of Sciences, Beijing, P.R. China), Jian Liang (CRIPAC & NLPR, CASIA; University of Chinese Academy of Sciences, Beijing, P.R. China), Haiqing Li (CRIPAC & NLPR, CASIA; University of Chinese Academy of Sciences, Beijing, P.R. China), and Zhenan Sun (CRIPAC & NLPR, CASIA; University of Chinese Academy of Sciences, Beijing, P.R. China; Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Beijing, P.R. China)

Every Smile is Unique: Landmark-Guided Diverse Smile Generation 7083
Wei Wang (University of Trento, Italy; Computer Vision Laboratory, EPFL, Lausanne, Switzerland), Xavier Alameda-Pineda (Inria, Grenoble-Alpes University, France), Dan Xu (University of Trento, Italy), Pascal Fua (Computer Vision Laboratory, EPFL, Lausanne, Switzerland), Elisa Ricci (University of Trento, Italy; Fondazione Bruno Kessler (FBK), Italy), and Nicu Sebe (University of Trento, Italy)

UV-GAN: Adversarial Facial UV Map Completion for Pose-Invariant Face Recognition 7093
Jiankang Deng (Imperial College London), Shiyang Cheng (Imperial College London), Niannan Xue (Imperial College London), Yuxiang Zhou (Imperial College London), and Stefanos Zafeiriou (Imperial College London)

Cascaded Pyramid Network for Multi-person Pose Estimation 7103
Yilun Chen (Tsinghua University, China), Zhicheng Wang (Tsinghua University, China), Yuxiang Peng (Tsinghua University, China), Zhiqiang Zhang (HuaZhong University of Science and Technology Megvii Inc. (Face++)), Gang Yu (Tsinghua University, China), and Jian Sun (Tsinghua University, China)

A Face-to-Face Neural Conversation Model 7113
Hang Chu (University of Toronto; Vector Institute), Daiqing Li (University of Toronto), Sanja Fidler (University of Toronto), and Sanja Fidler (Vector Institute)

End-to-End Recovery of Human Shape and Pose 7122
Angjoo Kanazawa (University of California, Berkeley), Michael J. Black (MPI for Intelligent Systems, Tübingen, Germany), David W. Jacobs (University of Maryland, College Park), and Jitendra Malik (University of California- Berkeley)

Oral and Spotlight 3-1A

O3-1A: Object Recognition & Scene Understanding IV

Squeeze-and-Excitation Networks 7132
Jie Hu (Momenta), Li Shen (University of Oxford), and Gang Sun (Momenta)

Revisiting Salient Object Detection: Simultaneous Detection, Ranking, and Subitizing of Multiple Salient Objects 7142
Md Amirul Islam (University of Manitoba), Mahmoud Kalash (University of Manitoba), and Neil D. B. Bruce (Ryerson University)

Context Encoding for Semantic Segmentation 7151
Hang Zhang (Rutgers University; Amazon Inc), Kristin Dana (Rutgers University), Jianping Shi (SenseTime), Zhongyue Zhang (Amazon Inc), Xiaogang Wang (The Chinese University of Hong Kong), Ambrish Tyagi (Amazon Inc), and Amit Agrawal (Amazon Inc)

S3-1A

Creating Capsule Wardrobes from Fashion Images 7161
Wei-Lin Hsiao (UT-Austin) and Kristen Grauman (UT-Austin)

Webly Supervised Learning Meets Zero-shot Learning: A Hybrid Approach for Fine-Grained Classification 7171
Li Niu (Rice University), Ashok Veeraraghavan (Rice University), and Ashu Sabharwal (Rice University)

Look, Imagine and Match: Improving Textual-Visual Cross-Modal Retrieval with Generative Models 7181
Jiuxiang Gu (Nanyang Technological University), Jianfei Cai (Nanyang Technological University), Shafiq Joty (Nanyang Technological University), Li Niu (Rice University, USA), and Gang Wang (Alibaba AI Labs, Hangzhou)

Bidirectional Attentive Fusion with Context Gating for Dense Video Captioning 7190
Jingwen Wang (South China University of Technology), Wenhao Jiang (Tencent AI Lab), Lin Ma (Tencent AI Lab), Wei Liu (Tencent AI Lab), and Yong Xu (South China University of Technology)

InLoc: Indoor Visual Localization with Dense Matching and View Synthesis 7199
Hajime Taira (Tokyo Institute of Technology), Masatoshi Okutomi (Tokyo Institute of Technology), Torsten Sattler (ETH Zurich), Mircea Cimpoi (CIIRC, CTU in Prague), Marc Pollefeys (ETH Zurich; Microsoft, Redmond), Josef Sivic (CIIRC, CTU in Prague; Inria), Tomas Pajdla (CIIRC, CTU in Prague), and Akihiko Torii (Tokyo Institute of Technology)

Towards High Performance Video Object Detection 7210
Xizhou Zhu (University of Science and Technology of China; Microsoft Research), Jifeng Dai (Microsoft Research), Lu Yuan (Microsoft Research), and Yichen Wei (Microsoft Research)

Neural Baby Talk 7219
Jiasen Lu (Georgia Institute of Technology), Jianwei Yang (Georgia Institute of Technology), Dhruv Batra (Georgia Institute of Technology; Facebook AI Research), and Devi Parikh (Georgia Institute of Technology; Facebook AI Research)

Few-Shot Image Recognition by Predicting Parameters from Activations 7229
Siyuan Qiao (Johns Hopkins University), Chenxi Liu (Johns Hopkins University), Wei Shen (Johns Hopkins University; Shanghai University), and Alan Yuille (Johns Hopkins University)

Iterative Visual Reasoning Beyond Convolutions 7239
Xinlei Chen (Carnegie Mellon University), Li-Jia Li (Google), Li Fei-Fei (Google), and Abhinav Gupta (Carnegie Mellon University)

Visual Question Reasoning on General Dependency Tree 7249
Qingxing Cao (Sun Yat-sen University, China), Xiaodan Liang (Sun Yat-sen University, China), Bailing Li (Sun Yat-sen University, China), Guanbin Li (Sun Yat-sen University, China), and Liang Lin (Sun Yat-sen University, China)

CVM-Net: Cross-View Matching Network for Image-Based Ground-to-Aerial Geo-Localization 7258
Sixing Hu (National University of Singapore), Mengdan Feng (National University of Singapore), Rang M. H. Nguyen (National University of Singapore), and Gim Hee Lee (National University of Singapore)

Revisiting Dilated Convolution: A Simple Approach for Weakly- and Semi-Supervised Semantic Segmentation 7268
Yunchao Wei (UIUC), Huaxin Xiao (NUS), Honghui Shi (IBM Thomas J. Watson Research Center), Zequn Jie (Tencent AI Lab), Jiashi Feng (NUS), and Thomas S. Huang (UIUC)

Low-Shot Learning from Imaginary Data 7278
Yu-Xiong Wang (Facebook AI Research (FAIR); Camegie Mellon University), Ross Girshick (Facebook AI Research (FAIR)), Martial Hebert (Camegie Mellon University), and Bharath Hariharan (Facebook AI Research (FAIR); Cornell University)

Oral and Spotlight 3-1B

O3-1B: Analyzing Humans

DoubleFusion: Real-Time Capture of Human Performances with Inner Body Shapes from a Single Depth Sensor 7287
Tao Yu (Tsinghua University, Beijing, China; Beihang University, Beijing, China), Zerong Zheng (Tsinghua University, Beijing, China), Kaiwen Guo (Tsinghua University, Beijing, China; Google Inc), Jianhui Zhao (Beihang University, Beijing, China), Qionghai Dai (Tsinghua University, Beijing, China), Hao Li (University of Southern California / USC Institute for Creative Technologies), Gerard Pons-Moll (Max-Planck-Institute for Informatics, Saarland Informatics Campus), and Yebin Liu (Tsinghua University, Beijing, China; Beijing National Research Center for Information Science and Technology (BNRist))

DensePose: Dense Human Pose Estimation in the Wild 7297
Riza Alp Güler (INRIA-CentraleSupélec), Natalia Neverova (Facebook AI Research), and Iasonas Kokkinos (Facebook AI Research)

Ordinal Depth Supervision for 3D Human Pose Estimation 7307
Georgios Pavlakos (University of Pennsylvania), Xiaowei Zhou (State Key Lab of CAD&CG, Zhejiang University), and Kostas Daniilidis (University of Pennsylvania)

S3-1B

Consensus Maximization for Semantic Region Correspondences 7317
Pablo Speciale (ETH Zurich), Danda P. Paude (Computer Vision Laboratory, D-ITET, ETH Zurich; ETH Zurich), Hayko Riemenschneider (Computer Vision Laboratory, D-ITET, ETH Zurich), Luc V. Gool (Computer Vision Laboratory, D-ITET, ETH Zurich; VISICS, ESAT/PSI, KU Leuven, Belgium), and Marc Pollefeys (ETH Zurich; Microsoft, Redmond, USA)

Robust Hough Transform Based 3D Reconstruction from Circular Light Fields 7327
Alessandro Vianello (Robert Bosch GmbH Robert Bosch Campus 1, Renningen, Germany), Jens Ackermann (Robert Bosch GmbH Robert Bosch Campus 1, Renningen, Germany), Maximilian Diebold (Heidelberg Collaboratory for Image Processing Berliner, Heidelberg, Germany), and Bernd Jähne (Heidelberg Collaboratory for Image Processing Berliner, Heidelberg, Germany)

Alive Caricature from 2D to 3D 7336
Qianyi Wu (University of Science and Technology of China), Juyong Zhang (University of Science and Technology of China), Yu-Kun Lai (Cardiff University), Jianmin Zheng (Nanyang Technological University, Singapore), and Jianfei Cai (Nanyang Technological University, Singapore)

Nonlinear 3D Face Morphable Model 7346
Luan Tran (Michigan State University) and Xiaoming Liu (Michigan State University)

Through-Wall Human Pose Estimation Using Radio Signals 7356
Mingmin Zhao (MIT CSAIL), Tianhong Li (MIT CSAIL), Mohammad Abu Alsheikh (MIT CSAIL), Yonglong Tian (MIT CSAIL), Hang Zhao (MIT CSAIL), Antonio Torralba (MIT CSAIL), and Dina Katabi (MIT CSAIL)

What Makes a Video a Video: Analyzing Temporal Information in Video Understanding Models and Datasets 7366
De-An Huang (Stanford University), Vignesh Ramanathan (Facebook), Dhruv Mahajan (Facebook), Lorenzo Torresani (Facebook; Dartmouth College), Manohar Paluri (Facebook), Li Fei-Fei (Stanford University), and Juan Carlos Niebles (Stanford University)

Fast Video Object Segmentation by Reference-Guided Mask Propagation 7376
Seoung Wug Oh (Yonsei University), Joon-Young Lee (Adobe Research), Kalyan Sunkavalli (Adobe Research), and Seon Joo Kim (Yonsei University)

NeuralNetwork-Viterbi: A Framework for Weakly Supervised Video Learning 7386
Alexander Richard (University of Bonn, Germany), Hilde Kuehne (University of Bonn, Germany), Ahsan Iqbal (University of Bonn, Germany), and Juergen Gall (University of Bonn, Germany)

Actor and Observer: Joint Modeling of First and Third-Person Videos 7396
Gunnar A. Sigurdsson (Carnegie Mellon University), Abhinav Gupta (Carnegie Mellon University), Cordelia Schmid (Inria), Ali Farhadi (Allen Institute for Artificial Intelligence), and Karteek Alahari (Inria)

HSA-RNN: Hierarchical Structure-Adaptive RNN for Video Summarization 7405
Bin Zhao (School of Computer Science and Center for OPTical IMagery Analysis and Learning (OPTIMAL), Northwestern Polytechnical University, Xi'an, Shaanxi, P. R. China), Xuelong Li (Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an, Shaanxi, P. R. China), and Xiaoqiang Lu (Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an, Shaanxi, P. R. China)

Fast and Accurate Online Video Object Segmentation via Tracking Parts 7415
Jingchun Cheng (Tsinghua University, China; University of California, Merced), Yi-Hsuan Tsai (NEC Laboratories America), Wei-Chih Hung (University of California, Merced), Shengjin Wang (Tsinghua University, China), and Ming-Hsuan Yang (University of California, Merced)

Now You Shake Me: Towards Automatic 4D Cinema 7425
Yuhao Zhoul (University of Toronto), Makarand Tapaswi (University of Toronto; Vector Institute), and Sanja Fidler (University of Toronto; Vector Institute)

Viewpoint-Aware Video Summarization 7435
Atsushi Kanehira (The University of Tokyo), Luc Van Gool (ETH Zurich; KU Leuven), Yoshitaka Ushiku (The University of Tokyo), and Tatsuya Harada (The University of Tokyo; RIKEN)

Oral and Spotlight 3-1C

O3-1C: Applications

Photometric Stereo in Participating Media Considering Shape-Dependent Forward Scatter 7445
Yuki Fujimura (Kyoto University, Japan), Masaaki Iiyama (Kyoto University, Japan), Atsushi Hashimoto (Kyoto University, Japan; OMRON SINIC X Corp., Japan), and Michihiko Minoh (Kyoto University, Japan)

Direction-Aware Spatial Context Features for Shadow Detection 7454
Xiaowei Hu (The Chinese University of Hong Kong), Lei Zhu (School of Nursing, The Hong Kong Polytechnic University), Chi-Wing Fu (The Chinese University of Hong Kong; Chinese Academy of Sciences, China), Jing Qin (School of Nursing, The Hong Kong Polytechnic University), and Pheng-Ann Heng (The Chinese University of Hong Kong; Chinese Academy of Sciences, China)

Discriminative Learning of Latent Features for Zero-Shot Recognition 7463
Yan Li (CRIPAC & NLPR, CASIA; University of Chinese Academy of Sciences), Junge Zhang (CRIPAC & NLPR, CASIA; University of Chinese Academy of Sciences), Jianguo Zhang (University of Dundee), and Kaiqi Huang (CRIPAC & NLPR, CASIA; University of Chinese Academy of Sciences; CAS Center for Excellence in Brain Science and Intelligence Technology)

S3-1C

Learning to Adapt Structured Output Space for Semantic Segmentation 7472
Yi-Hsuan Tsai (NEC Laboratories America), Wei-Chih Hung (University of California, Merced), Samuel Schulter (NEC Laboratories America), Kihyuk Sohn (NEC Laboratories America), Ming-Hsuan Yang (University of California, Merced), and Manmohan Chandraker (NEC Laboratories America; University of California, San Diego)

Multi-task Learning Using Uncertainty to Weigh Losses for Scene Geometry and Semantics 7482
Roberto Cipolla (University of Cambridge), Yarin Gal (University of Cambridge), and Alex Kendall (University of Cambridge)

Jointly Localizing and Describing Events for Dense Video Captioning 7492
Yehao Li (Sun Yat-sen University, China), Ting Yao (Microsoft Research, Beijing), Yingwei Pan (University of Science and Technology of China), Hongyang Chao (Sun Yat-sen University, China), and Tao Mei (Microsoft Research, Beijing)

Going from Image to Video Saliency: Augmenting Image Salience with Dynamic Attentional Push 7501
Siavash Gorji (McGill University Montreal, Quebec, Canada) and James J. Clark (McGill University Montreal, Quebec, Canada)

M3: Multimodal Memory Modelling for Video Captioning 7512
Junbo Wang ((CRIPAC); (UCAS)), Wei Wang ((CRIPAC); (UCAS)), Yan Huang ((CRIPAC); (UCAS)), Liang Wang ((CRIPAC); (UCAS)), and Tieniu Tan ((CRIPAC); (UCAS))

Emotional Attention: A Study of Image Sentiment and Visual Attention 7521
Shaojing Fan (National University of Singapore), Zhiqi Shen (National University of Singapore), Ming Jiang (University of Minnesota), Bryan L. Koenig (Southern Utah University), Juan Xu (University of Minnesota), Mohan S. Kankanhalli (National University of Singapore), and Qi Zhao (University of Minnesota)

A Low Power, High Throughput, Fully Event-Based Stereo System 7532
Alexander Andreopoulos (IBM Research), Hirak J. Kashyap (IBM Research), Tapan K. Nayak (IBM Research), Arnon Amir (IBM Research), and Myron D. Flickner (IBM Research)

VITON: An Image-Based Virtual Try-on Network 7543
Xintong Han (University of Maryland, College Park), Zuxuan Wu (University of Maryland, College Park), Zhe Wu (University of Maryland, College Park), Ruichi Yu (University of Maryland, College Park), and Larry S. Davis (University of Maryland, College Park)

Multi-oriented Scene Text Detection via Corner Localization and Region Segmentation 7553
Pengyuan Lyu (Huazhong University of Science and Technology; Megvii (Face++) Technology Inc.), Cong Yao (Megvii (Face++) Technology Inc.), Wenhao Wu (Megvii (Face++) Technology Inc.), Shuicheng Yan (National University of Singapore), and Xiang Bai (Huazhong University of Science and Technology)

Multi-content GAN for Few-Shot Font Style Transfer 7564
Samaneh Azadi (UC Berkeley), Matthew Fisher (Adobe Research), Vladimir Kim (Adobe Research), Zhaowen Wang (Adobe Research), Eli Shechtman (Adobe Research), and Trevor Darrell (UC Berkeley)

Audio to Body Dynamics 7574
Eli Shlizerman (Facebook Inc.; University of Washington), Lucio Dery (Facebook Inc.; Stanford University), Hayden Schoen (Facebook Inc.), Ira Kemelmacher-Shlizerman (Facebook Inc.), and Ira Kemelmacher-Shlizerman (University of Washington)

Weakly Supervised Coupled Networks for Visual Sentiment Analysis 7584
Jufeng Yang (Nankai University, Tianjin, China), Dongyu She (Nankai University, Tianjin, China), Yu-Kun Lai (Cardiff University, Cardiff), Paul L. Rosin (Cardiff University, Cardiff), and Ming-Hsuan Yang (University of California, Merced, USA)

Future Person Localization in First-Person Videos 7593
Takuma Yagi (The University of Tokyo Tokyo, Japan), Karttikeya Mangalam (Indian Institute of Technology Kanpur, India), Ryo Yonetani (The University of Tokyo Tokyo, Japan), and Yoichi Sato (The University of Tokyo Tokyo, Japan)

Poster Session P3-1

Preserving Semantic Relations for Zero-Shot Learning 7603
Soma Biswas (Indian Institute of Science) and Yashas Annadani (National Institute of Technology - Karnataka)

Show Me a Story: Towards Coherent Neural Story Illustration 7613
Hareesh Ravi (Rutgers University), Lezi Wang (Rutgers University), Carlos M. Muniz (Rutgers University), Leonid Sigal (University of British Columbia), Dimitris N. Metaxas (Rutgers University), and Mubbasir Kapadia (Rutgers University)

Reconstruction Network for Video Captioning 7622
Bairui Wang (Shandong University, China), Lin Ma (Tencent AI Lab), Wei Zhang (Shandong University, China), and Wei Liu (Tencent AI Lab)

Fast Spectral Ranking for Similarity Search 7632
Ahmet Iscen (VRG, FEE, CTU in Prague), Yannis Avrithis (Inria Rennes), Giorgos Tolias (VRG, FEE, CTU in Prague), Teddy Furon (Inria Rennes), and Ondrej Chum (VRG, FEE, CTU in Prague)

Mining on Manifolds: Metric Learning Without Labels 7642
Ahmet Iscen (VRG, FEE, CTU in Prague), Giorgos Tolias (VRG, FEE, CTU in Prague), Yannis Avrithis (Inria Rennes), and Ondrej Chum (VRG, FEE, CTU in Prague)

PIXOR: Real-time 3D Object Detection from Point Clouds 7652
Bin Yang (Uber Advanced Technologies Group University of Toronto), Wenjie Luo (Uber Advanced Technologies Group University of Toronto), and Raquel Urtasun (Uber Advanced Technologies Group University of Toronto)

Leveraging Unlabeled Data for Crowd Counting by Learning to Rank 7661
Xialei Liu (Computer Vision Center Barcelona, Spain), Joost van de Weijer (Computer Vision Center Barcelona, Spain), and Andrew D. Bagdanov (MICC, University of Florence Florence, Italy)

Zero-Shot Kernel Learning 7670
Hongguang Zhang (Australian National University; Data61/CSIRO) and Piotr Koniusz (Data61/CSIRO; Australian National University)

Differential Attention for Visual Question Answering 7680
Badri Patro (IIT Kanpur) and Vinay P. Namboodiri (IIT Kanpur)

Learning from Noisy Web Data with Category-Level Supervision 7689
Li Niu (Rice University), Qingtao Tang (Tsinghua University, China), Ashok Veeraraghavan (Rice University), and Ashu Sabharwal (Rice University)

Toward Driving Scene Understanding: A Dataset for Learning Driver Behavior and Causal Reasoning 7699
Vasili Ramanishka (Boston University), Yi-Ting Chen (Honda Research Institute, USA), Teruhisa Misu (Honda Research Institute, USA), and Kate Saenko (Boston University)

Learning Attribute Representations with Localization for Flexible Fashion Search 7708
*Kenan E. Ak (National University of Singapore, Singapore; Institute for Infocomm Research, A*STAR, Singapore), Ashraf A. Kassim (National University of Singapore, Singapore), Joo Hwee Lim (Institute for Infocomm Research, A*STAR, Singapore), and Jo Yew Tham (ESP xMedia Pte. Ltd., Singapore)*

Bidirectional Retrieval Made Simple 7718
Jônatas Wehrmann (Pontifícia Universidade Católica do Rio Grande do Sul) and Rodrigo C. Barros (Pontifícia Universidade Católica do Rio Grande do Sul)

Learning Multi-instance Enriched Image Representations via Non-greedy Ratio Maximization of the l1-Norm Distances 7727
Kai Liu (Colorado School of Mines, Golden, Colorado, USA), Hua Wang (Colorado School of Mines, Golden, Colorado, USA), Feiping Nie (Northwestern Polytechnical University, Xi'an, P. R. China), and Hao Zhang (Colorado School of Mines, Golden, Colorado, USA)

Learning Visual Knowledge Memory Networks for Visual Question Answering 7736
Zhou Su (Tencent Wechat), Chen Zhu (ShanghaiTech University), Yinpeng Dong (Tsinghua University, China), Dongqi Cai (Intel Labs China), Yurong Chen (Intel Labs China), and Jianguo Li (Intel Labs China)

Visual Grounding via Accumulated Attention 7746
Chaorui Deng (South China University of Technology, China), Qi Wu (Australia Centre for Robotic Vision, The University of Adelaide, Australia), Qingyao Wu (South China University of Technology, China), Fuyuan Hu (Suzhou University of Science and Technology, China), Fan Lyu (Suzhou University of Science and Technology, China), and Mingkui Tan (South China University of Technology, China)

Beyond Trade-Off: Accelerate FCN-Based Face Detector with Higher Accuracy 7756
Guanglu Song (Beihang University), Yu Liu (The Chinese University of Hong Kong), Ming Jiang (Beihang University), Yujie Wang (Beihang University), Junjie Yan (Sensetime Group Limited), and Biao Leng (Beihang University)

PackNet: Adding Multiple Tasks to a Single Network by Iterative Pruning 7765
Arun Mallya (University of Illinois at Urbana-Champaign) and Svetlana Lazebnik (University of Illinois at Urbana-Champaign)

Repulsion Loss: Detecting Pedestrians in a Crowd 7774
Xinlong Wang (Tongji University), Tete Xiao (Peking University), Yuning Jiang (Megvii, Inc.), Shuai Shao (Megvii, Inc.), Jian Sun (Megvii, Inc.), and Chunhua Shen (The University of Adelaide)

Neural Sign Language Translation 7784
Necati Cihan Camgoz (University of Surrey), Simon Hadfield (University of Surrey), Oscar Koller (RWTH Aachen University), Hermann Ney (RWTH Aachen University), and Richard Bowden (University of Surrey)

Non-local Neural Networks 7794
Xiaolong Wang (Carnegie Mellon University; Facebook AI Research), Ross Girshick (Facebook AI Research), Abhinav Gupta (Carnegie Mellon University), and Kaiming He (Facebook AI Research)

LAMV: Learning to Align and Match Videos with Kernelized Temporal Layers 7804
Lorenzo Baraldi (University of Modena and Reggio Emilia), Matthijs Douze (Facebook AI Research), Rita Cucchiara (University of Modena and Reggio Emilia), and Herve Jegou (Facebook AI Research)

Optimizing Video Object Detection via a Scale-Time Lattice 7814
Kai Chen (CUHK - SenseTime Joint Lab, The Chinese University of Hong Kong), Jiaqi Wang (CUHK - SenseTime Joint Lab, The Chinese University of Hong Kong), Shuo Yang (CUHK - SenseTime Joint Lab, The Chinese University of Hong Kong; Amazon Rekognition), Xingcheng Zhang (CUHK - SenseTime Joint Lab, The Chinese University of Hong Kong), Yuanjun Xiong (CUHK - SenseTime Joint Lab, The Chinese University of Hong Kong; Amazon Rekognition), Chen Change Loy (CUHK - SenseTime Joint Lab, The Chinese University of Hong Kong), and Dahua Lin (CUHK - SenseTime Joint Lab, The Chinese University of Hong Kong)

Learning Compressible 360° Video Isomers 7824
Yu-Chuan Su (The University of Texas at Austin) and Kristen Grauman (The University of Texas at Austin)

Attention Clusters: Purely Attention Based Local Feature Integration for Video Classification 7834
Xiang Long (Tsinghua University, China), Chuang Gan (MIT-IBM Watson AI Lab), Gerard de Melo (Rutgers University), Jiajun Wu (MIT), Xiao Liu (Baidu IDL), and Shilei Wen (Baidu IDL)

What have We Learned from Deep Representations for Action Recognition? 7844
Axel Pinz (TU Graz), Andrew Zisserman (University of Oxford), Richard P. Wildes (York University, Toronto), and Christoph Feichtenhofer (TU Graz)

Controllable Video Generation with Sparse Trajectories 7854
Zekun Hao (Cornell University), Xun Huang (Cornell University), and Serge Belongie (Cornell University)

Representing and Learning High Dimensional Data with the Optimal Transport Map from a Probabilistic Viewpoint 7864
Serim Park (Carnegie Mellon University) and Matthew Thorpe (The University of Cambridge Cambridge)

CLIP-Q: Deep Network Compression Learning by In-parallel Pruning-Quantization 7873
Frederick Tung (Simon Fraser University) and Greg Mori (Simon Fraser University)

Inference in Higher Order MRF-MAP Problems with Small and Large Cliques 7883
Ishant Shanu (IIIT Delhi), Chetan Arora (IIIT Delhi), and S.N. Maheshwari (IIT Delhi)

ROAD: Reality Oriented Adaptation for Semantic Segmentation of Urban Scenes 7892
Yuhua Chen (Computer Vision Laboratory, ETH Zurich), Wen Li (Computer Vision Laboratory, ETH Zurich), and Luc Van Gool (Computer Vision Laboratory, ETH Zurich; VISICS, ESAT/PSI, KU Leuven)

Eye In-painting with Exemplar Generative Adversarial Networks 7902
Brian Dolhansky (Facebook Inc.) and Cristian Canton Ferrer (Facebook Inc.)

clcNet: Improving the Efficiency of Convolutional Neural Network Using Channel Local Convolutions 7912
Dong-Qing Zhang (ImaginationAI LLC)

Towards Effective Low-Bitwidth Convolutional Neural Networks 7920
Bohan Zhuang (The University of Adelaide, Australia; Australian Centre for Robotic Vision), Chunhua Shen (The University of Adelaide, Australia; Australian Centre for Robotic Vision), Mingkui Tan (South China University of Technology, China), Lingqiao Liu (The University of Adelaide, Australia), and Ian Reid (The University of Adelaide, Australia; Australian Centre for Robotic Vision)

Stochastic Downsampling for Cost-Adjustable Inference and Improved Regularization in Convolutional Networks 7929
Jason Kuen (Nanyang Technological University), Xiangfei Kong (Nanyang Technological University), Zhe Lin (Adobe Research), Gang Wang (Alibaba AI Labs), Jianxiong Yin (NVIDIA), Simon See (NVIDIA), and Yap-Peng Tan (Nanyang Technological University)

Face Aging with Identity-Preserved Conditional Generative Adversarial Networks 7939
Xu Tang (Baidu), Zongwei Wang (Shanghaitech University), Weixin Luo (Shanghaitech University), and Shenghua Gao (Shanghaitech University)

Unsupervised Cross-Dataset Person Re-identification by Transfer Learning of Spatial-Temporal Patterns 7948
Jianming Lv (South China University of Technology), Weihang Chen (South China University of Technology), Qing Li (City University of Hongkongy), and Can Yang (South China University of Technology)

Feature Quantization for Defending Against Distortion of Images 7957
Zhun Sun (Tohoku University; RIKEN Center for AIP), Mete Ozay (Tohoku University), Yan Zhang (RIKEN Center for AIP), Xing Liu (Tohoku University), and Takayuki Okatani (Tohoku University; RIKEN Center for AIP)

Tagging Like Humans: Diverse and Distinct Image Annotation 7967
Baoyuan Wu (Tencent AI Lab), Weidong Chen (Tencent AI Lab), Peng Sun (Tencent AI Lab), Wei Liu (Tencent AI Lab), Bernard Ghanem (KAUST), and Siwei Lyu (University at Albany, SUNY)

Re-weighted Adversarial Adaptation Network for Unsupervised Domain Adaptation 7976
Qingchao Chen (University College London), Yang Liu (University of Cambridge; Adobe Research), Zhaowen Wang (Adobe Research), Ian Wassell (University of Cambridge; Adobe Research), and Kevin Chetty (University College London)

Inferring Semantic Layout for Hierarchical Text-to-Image Synthesis 7986
Seunghoon Hong (University of Michigan), Dingdong Yang (University of Michigan), Jongwook Choi (University of Michigan), and Honglak Lee (University of Michigan; Google Brain)

Regularizing RNNs for Caption Generation by Reconstructing the Past with the Present 7995
Xinpeng Chen (Wuhan University), Lin Ma (Tencent AI Lab), Wenhao Jiang (Tencent AI Lab), Jian Yao (Wuhan University), and Wei Liu (Tencent AI Lab)

Unsupervised Domain Adaptation with Similarity Learning 8004
Pedro O. Pinheiro (Element AI Montreal, QC, Canada)

Learning Deep Sketch Abstraction 8014
Umar Riaz Muhammad (SketchX, Queen Mary University of London), Yongxin Yang (The University of Edinburgh), Yi-Zhe Song (SketchX, Queen Mary University of London), Tao Xiang (SketchX, Queen Mary University of London), and Timothy M. Hospedales (SketchX, Queen Mary University of London; The University of Edinburgh)

Matching Adversarial Networks 8024
Gellert Mattyus (Uber Advanced Technologies Group and University of Toronto) and Raquel Urtasun (Uber Advanced Technologies Group and University of Toronto)

SoS-RSC: A Sum-of-Squares Polynomial Approach to Robustifying Subspace Clustering Algorithms 8033
Mario Sznaier (Northeastern University, Boston, MA) and Octavia Camps (Northeastern University, Boston, MA)

Resource Aware Person Re-identification Across Multiple Resolutions 8042
Yan Wang (Cornell University), Lequn Wang (Cornell University), Yurong You (Shanghai Jiao Tong University), Xu Zou (Tsinghua University, China), Vincent Chen (Cornell University), Serena Li (Cornell University), Gao Huang (Cornell University), Bharath Hariharan (Cornell University), and Kilian Q. Weinberger (Cornell University)

Learning and Using the Arrow of Time 8052
Donglai Wei (Harvard University), Joseph Lim (University of Southern California), Andrew Zisserman (University of Oxford), and William T Freeman (Massachusetts Institute of Technology; Google Research)

Neural Style Transfer via Meta Networks 8061
Falong Shen (Peking University), Shuicheng Yan (360 AI Institute; National University of Singapore), and Gang Zeng (Peking University)

People, Penguins and Petri Dishes: Adapting Object Counting Models to New Visual Domains and Object Types Without Forgetting 8070
Mark Marsden (Insight Centre for Data Analytics Dublin City University, Ireland), Kevin McGuinness (Insight Centre for Data Analytics Dublin City University, Ireland), Suzanne Little (Insight Centre for Data Analytics Dublin City University, Ireland), Ciara E. Keogh (Conway Institute, School of Medicine University College Dublin, Ireland), and Noel E. O'Connor (Insight Centre for Data Analytics Dublin City University, Ireland)

HydraNets: Specialized Dynamic Architectures for Efficient Inference 8080
Noam Shazeer (Google Inc.), Kayvon Fatahalian (Stanford University), William R.Mark (Google Inc.), and Ravi Teja Mullapudi (CMU)

SketchMate: Deep Hashing for Million-Scale Human Sketch Retrieval 8090
Peng Xu (Beijing University of Posts and Telecommunications, China; SketchX, Queen Mary University of London; Beijing University of Posts and Telecommunications, China; Queen Mary University of London), Yongye Huang (Beijing University of Posts and Telecommunications, China), Tongtong Yuan (Beijing University of Posts and Telecommunications, China), Kaiyue Pang (SketchX, Queen Mary University of London), Yi-Zhe Song (SketchX, Queen Mary University of London), Tao Xiang (SketchX, Queen Mary University of London), Timothy M. Hospedales (The University of Edinburgh), Zhanyu Ma (Beijing University of Posts and Telecommunications, China), and Jun Guo (Beijing University of Posts and Telecommunications, China)

From Source to Target and Back: Symmetric Bi-Directional Adaptive GAN 8099
Paolo Russo (Sapienza University of Rome, Italy; Italian Institute of Technology), Fabio M. Carlucci (Sapienza University of Rome, Italy; Italian Institute of Technology), Tatiana Tommasi (Italian Institute of Technology), and Barbara Caputo (Sapienza University of Rome, Italy; Italian Institute of Technology)

OLE: Orthogonal Low-rank Embedding, A Plug and Play Geometric Loss for Deep Learning 8109
Jose Lezama (Universidad de la Republica, Uruguay), Qiang Qiu (Duke University, USA), Pablo Musé (Universidad de la Republica, Uruguay), and Guillermo Sapiro (Duke University, USA)

Efficient Parametrization of Multi-domain Deep Neural Networks 8119
Sylvestre-Alvise Rebuffi (Visual Geometry Group University of Oxford), Andrea Vedaldi (Visual Geometry Group University of Oxford), and Hakan Bilen (School of Informatics University of Edinburgh)

Deep Density Clustering of Unconstrained Faces 8128
Wei-An Lin (University of Maryland, College Park), Jun-Cheng Chen (University of Maryland, College Park), Carlos D. Castillo (University of Maryland, College Park), and Rama Chellappa (University of Maryland, College Park)

Geometric Multi-model Fitting with a Convex Relaxation Algorithm 8138
Paul Amayo (Oxford Robotics Institute University of Oxford), Pedro Pinies (Oxford Robotics Institute University of Oxford), Lina M. Paz (Oxford Robotics Institute University of Oxford), and Paul Newman (Oxford Robotics Institute University of Oxford)

Fast and Robust Estimation for Unit-Norm Constrained Linear Fitting Problems 8147
Daiki Ikami (The University of Tokyo, Japan), Toshihiko Yamasaki (The University of Tokyo, Japan), and Kiyoharu Aizawa (The University of Tokyo, Japan)

Importance Weighted Adversarial Nets for Partial Domain Adaptation 8156
Jing Zhang (Advanced Multimedia Research Lab, University of Wollongong, Australia), Zewei Ding (Advanced Multimedia Research Lab, University of Wollongong, Australia), Wanqing Li (Advanced Multimedia Research Lab, University of Wollongong, Australia), and Philip Ogunbona (Advanced Multimedia Research Lab, University of Wollongong, Australia)

Efficient Subpixel Refinement with Symbolic Linear Predictors 8165
Vincent Lui (Monash University, Victoria, Australia), Jonathon Geeves (Monash University, Victoria, Australia), Winston Yii (Monash University, Victoria, Australia), and Tom Drummond (Monash University, Victoria, Australia)

Scale-Recurrent Network for Deep Image Deblurring 8174
Xin Tao (The Chinese University of Hong Kong; YouTu Lab, Tencent), Hongyun Gao (The Chinese University of Hong Kong; YouTu Lab, Tencent), Xiaoyong Shen (YouTu Lab, Tencent), Jue Wang (Megvii Inc.), and Jiaya Jia (The Chinese University of Hong Kong; YouTu Lab, Tencent)

DeblurGAN: Blind Motion Deblurring Using Conditional Adversarial Networks 8183
Orest Kupyn (Ukrainian Catholic University, Lviv, Ukraine; ELEKS Ltd.), Volodymyr Budzan (Ukrainian Catholic University, Lviv, Ukraine; ELEKS Ltd.), Mykola Mykhailych (Ukrainian Catholic University, Lviv, Ukraine), Dmytro Mishkin (Visual Recognition Group, Center for Machine Perception, FEE, CTU in Prague), and Jiri Matas (Visual Recognition Group, Center for Machine Perception, FEE, CTU in Prague)

A2-RL: Aesthetics Aware Reinforcement Learning for Image Cropping 8193
Debang Li (CRIPAC & NLPR, Institute of Automation, Chinese Academy of Sciences, Beijing, China; University of Chinese Academy of Sciences, Beijing, China), Huikai Wu (CRIPAC & NLPR, Institute of Automation, Chinese Academy of Sciences, Beijing, China; University of Chinese Academy of Sciences, Beijing, China), Junge Zhang (CRIPAC & NLPR, Institute of Automation, Chinese Academy of Sciences, Beijing, China; University of Chinese Academy of Sciences, Beijing, China), and Kaiqi Huang (CRIPAC & NLPR, Institute of Automation, Chinese Academy of Sciences, Beijing, China; University of Chinese Academy of Sciences, Beijing, China; CAS Center for Excellence in Brain Science and Intelligence Technology, Beijing, China)

Single Image Dehazing via Conditional Generative Adversarial Network 8202
Runde Li (Nanjing University of Science and Technology, China), Jinshan Pan (Nanjing University of Science and Technology, China), Zechao Li (Nanjing University of Science and Technology, China), and Jinhui Tang (Nanjing University of Science and Technology, China)

On the Duality Between Retinex and Image Dehazing 8212
Adrian Galdran (INESC TEC Porto, Portugal), Alessandro Bria (University of Cassino and L.M., Italy), Aitor Alvarez-Gila (Tecnalia, Spain Computer Vision Center, Spain), Javier Vazquez-Corral (Universitat Pompeu Fabra, Spain), and Marcelo Bertalmí-o (Universitat Pompeu Fabra, Spain)

Arbitrary Style Transfer with Deep Feature Reshuffle 8222
Shuyang Gu (University of Science and Technology of China), Congliang Chen (Peking University), Jing Liao (Microsoft Research), and Lu Yuan (Microsoft Research)

Nonlocal Low-Rank Tensor Factor Analysis for Image Restoration 8232
Xinyuan Zhang (Duke University Durham, NC, USA), Xin Yuan (Nokia Bell Labs Murray Hill, NJ, USA), and Lawrence Carin (Duke University Durham, NC, USA)

Avatar-Net: Multi-scale Zero-Shot Style Transfer by Feature Decoration 8242
Lu Sheng (CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong), Ziyi Lin (SenseTime Research), Jing Shao (SenseTime Research), and Xiaogang Wang (CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong)

Missing Slice Recovery for Tensors Using a Low-Rank Model in Embedded Space 8251
Tatsuya Yokota (Nagoya Institute of Technology, Gokiso, Showa, Nagoya, Japan), Burak Erem (Boston Children's Hospital, Boston, MA, USA), Seyhmus Guler (Boston Children's Hospital, Boston, MA, USA), Simon K. Warfield (Boston Children's Hospital, Boston, MA, USA), and Hidekata Hontani (Nagoya Institute of Technology, Gokiso, Showa, Nagoya, Japan)

Deep Semantic Face Deblurring 8260
Ziyi Shen (Beijing Institute of Technology), Wei-Sheng Lai (University of California, Merced), Tingfa Xu (Beijing Institute of Technology), Jan Kautz (Nvidia), and Ming-Hsuan Yang (University of California, Merced; Google Cloud)

GraphBit: Bitwise Interaction Mining via Deep Reinforcement Learning 8270
Yueqi Duan (Tsinghua University, China; State Key Lab of Intelligent Technologies and Systems; Beijing National Research Center for Information Science and Technology), Ziwei Wang (Tsinghua University, China), Jiwen Lu (Tsinghua University, China; State Key Lab of Intelligent Technologies and Systems; Beijing National Research Center for Information Science and Technology), Xudong Lin (Tsinghua University, China), and Jie Zhou (Tsinghua University, China; State Key Lab of Intelligent Technologies and Systems; Beijing National Research Center for Information Science and Technology)

Recurrent Saliency Transformation Network: Incorporating Multi-stage Visual Cues for Small Organ Segmentation 8280
Qihang Yu (Peking University), Lingxi Xie (The Johns Hopkins University), Yan Wang (The Johns Hopkins University), Yuyin Zhou (The Johns Hopkins University), Elliot K. Fishman (The Johns Hopkins Medical Institute), and Alan L. Yuille (The Johns Hopkins University)

Thoracic Disease Identification and Localization with Limited Supervision 8290
Zhe Li (Syracuse University), Chong Wang (Google Inc.), Mei Han (PingAn Technology, US Research Lab), Yuan Xue (Google Inc.), Wei Wei (Google Inc.), Li-Jia Li (Google Inc.), and Li Fei-Fei (Google Inc.)

Quantization of Fully Convolutional Networks for Accurate Biomedical Image Segmentation 8300
Xiaowei Xu (Huazhong University of Science and Technology), Qing Lu (Univerity of Notre Dame), Lin Yang (Univerity of Notre Dame), Sharon Hu (Univerity of Notre Dame), Danny Chen (Univerity of Notre Dame), Yu Hu (Huazhong University of Science and Technology), and Yiyu Shi (Univerity of Notre Dame)

Visual Feature Attribution Using Wasserstein GANs 8309
Christian F. Baumgartner (Computer Vision Lab, ETH Zurich), Lisa M. Koch (Computer Vision and Geometry Group, ETH Zurich), Kerem Can Tezcan (Computer Vision Lab, ETH Zurich), and Jia Xi Ang (Computer Vision Lab, ETH Zurich)

O3-2A: Analyzing Humans in Images II

Total Capture: A 3D Deformation Model for Tracking Faces, Hands, and Bodies 8320
Hanbyul Joo (Carnegie Mellon University), Tomas Simon (Carnegie Mellon University; Facebook Reality Labs, Pittsburgh), and Yaser Sheikh (Carnegie Mellon University; Facebook Reality Labs, Pittsburgh)

Augmented Skeleton Space Transfer for Depth-Based Hand Pose Estimation 8330
Seungryul Baek (Imperial College London), Kwang In Kim (University of Bath), and Tae-Kyun Kim (Imperial College London)

Synthesizing Images of Humans in Unseen Poses 8340
Guha Balakrishnan (MIT), Amy Zhao (MIT), Adrian V. Dalca (MIT and MGH), Fredo Durand (MIT), and John Guttag (MIT)

Oral and Spotlight 3-2A

SSNet: Scale Selection Network for Online 3D Action Prediction 8349
Jun Liu (Nanyang Technological University, Singapore), Amir Shahroudy (Nanyang Technological University, Singapore), Gang Wang (Alibaba Group, Hangzhou, China), Ling-Yu Duan (Peking University, Beijing, China), and Alex C. Kot (Nanyang Technological University, Singapore)

Detecting and Recognizing Human-Object Interactions 8359
Georgia Gkioxari (Facebook AI Research (FAIR)), Ross Girshick (Facebook AI Research (FAIR)), Piotr Dollár (Facebook AI Research (FAIR)), and Kaiming He (Facebook AI Research (FAIR))

Unsupervised Learning and Segmentation of Complex Activities from Video 8368
Fadime Sener (University of Bonn, Germany) and Angela Yao (University of Bonn, Germany)

Unsupervised Training for 3D Morphable Model Regression 8377
Kyle Genova (Princeton University; Google Research), Forrester Cole (Google Research), Aaron Maschinot (Google Research), Aaron Sarna (Google Research), Daniel Vlasic (Google Research), and William T. Freeman (Google Research; MIT CSAIL)

Video Based Reconstruction of 3D People Models 8387
Thiemo Alldieck (Computer Graphics Lab, TU Braunschweig, Germany), Marcus Magnor (Computer Graphics Lab, TU Braunschweig, Germany), Weipeng Xu (Max Planck Institute for Informatics, Saarland Informatics Campus, Germany), Christian Theobalt (Max Planck Institute for Informatics, Saarland Informatics Campus, Germany), and Gerard Pons-Moll (Max Planck Institute for Informatics, Saarland Informatics Campus, Germany)

Pose-Guided Photorealistic Face Rotation 8398
Yibo Hu (CRIPAC & NLPR & CEBSIT, CASIA), Yibo Hu (University of Chinese Academy of Sciences), Xiang Wu (CRIPAC & NLPR & CEBSIT, CASIA), Bing Yu (Noah's Ark Laboratory, Huawei Technologies Co., Ltd.), Ran He (CRIPAC & NLPR & CEBSIT, CASIA; University of Chinese Academy of Sciences), and Zhenan Sun (CRIPAC & NLPR & CEBSIT, CASIA; University of Chinese Academy of Sciences)

Mesoscopic Facial Geometry Inference Using Deep Neural Networks 8407
Loc Huynh (USC Institute for Creative Technologies), Weikai Chen (USC Institute for Creative Technologies), Shunsuke Saito (USC Institute for Creative Technologies; University of Southern California; Pinscreen), Jun Xing (USC Institute for Creative Technologies), Koki Nagano (Pinscreen), Andrew Jones (USC Institute for Creative Technologies), Paul Debevec (USC Institute for Creative Technologies), and Hao Li (USC Institute for Creative Technologies; University of Southern California; Pinscreen)

Hand PointNet: 3D Hand Pose Estimation Using Point Sets 8417
Liuhao Ge (Institute for Media Innovation, Interdisciplinary Graduate School, Nanyang Technological University), Yujun Cai (Institute for Media Innovation, Interdisciplinary Graduate School, Nanyang Technological University), Junwu Weng (Nanyang Technological University), and Junsong Yuan (State University of New York at Buffalo)

Seeing Voices and Hearing Faces: Cross-Modal Biometric Matching 8427
Arsha Nagrani (University of Oxford), Samuel Albanie (University of Oxford), and Andrew Zisserman (University of Oxford)

Learning Monocular 3D Human Pose Estimation from Multi-view Images 8437
Helge Rhodin (CVLab, EPFL, Lausanne, Switzerland), Frédéric Meyer (UNIL, Lausanne, Switzerland), Jörg Spörri (Balgrist University Hospital, Zurich, Switzerland; University of Salzburg, Salzburg, Austria), Erich Müller (University of Salzburg, Salzburg, Austria), Victor Constantin (CVLab, EPFL, Lausanne, Switzerland), Pascal Fua (CVLab, EPFL, Lausanne, Switzerland), Isinsu Katircioglu (CVLab, EPFL, Lausanne, Switzerland), and Mathieu Salzmann (CVLab, EPFL, Lausanne, Switzerland)

Separating Style and Content for Generalized Style Transfer 8447
Yexun Zhang (Shanghai Jiao Tong University), Ya Zhang (Shanghai Jiao Tong University), and Wenbin Cai (Microsoft)

TextureGAN: Controlling Deep Image Synthesis with Texture Patches 8456
Wenqi Xian (Georgia Institute of Technology), Patsorn Sangkloy (Georgia Institute of Technology), Varun Agrawal (Georgia Institute of Technology), Amit Raj (Georgia Institute of Technology), Jingwan Lu (Adobe Research), Chen Fang (Adobe Research), Fisher Yu (UC Berkeley), and James Hays (Georgia Institute of Technology; Argo AI)

Connecting Pixels to Privacy and Utility: Automatic Redaction of Private Information in Images 8466
Tribhuvanesh Orekondy (Max Planck Institute for Informatics Saarland Informatics Campus Saarbrücken, Germany), Mario Fritz (Max Planck Institute for Informatics Saarland Informatics Campus Saarbrücken, Germany), and Bernt Schiele (Max Planck Institute for Informatics Saarland Informatics Campus Saarbrücken, Germany)

O3-2B: Machine Learning for Computer Vision IV

MapNet: An Allocentric Spatial Memory for Mapping Environments 8476
Joao F. Henriques (Visual Geometry Group, University of Oxford) and Andrea Vedaldi (Visual Geometry Group, University of Oxford)

Accurate and Diverse Sampling of Sequences Based on a "Best of Many" Sample Objective 8485
Apratim Bhattacharyya (Max Planck Institute for Informatics, Saarland Informatics Campus, Saarbrücken, Germany), Bernt Schiele (Max Planck Institute for Informatics, Saarland Informatics Campus, Saarbrücken, Germany), and Mario Fritz (Max Planck Institute for Informatics, Saarland Informatics Campus, Saarbrücken, Germany)

VirtualHome: Simulating Household Activities Via Programs 8494
Xavier Puig (MIT), Kevin Ra (McGill University), Marko Boben (University of Ljubljana), Jiaman Li (University of Toronto), Tingwu Wang (University of Toronto), Sanja Fidler (University of Toronto), and Antonio Torralba (MIT)

S3-2B

Generate To Adapt: Aligning Domains Using Generative Adversarial Networks 8503
Swami Sankaranarayanan (University of Maryland), Yogesh Balaji (University of Maryland), Carlos D. Castillo (University of Maryland), and Rama Chellappa (University of Maryland)

Multi-agent Diverse Generative Adversarial Networks 8513
Arnab Ghosh (University of Oxford), Viveka Kulharia (University of Oxford), Vinay Namboodiri (IIT Kanpur, India), Philip H.S. Torr (University of Oxford), and Puneet K. Dokania (University of Oxford)

A PID Controller Approach for Stochastic Optimization of Deep Networks 8522
Wangpeng An (Tsinghua University, Shenzhen, China; The Hong Kong Polytechnic University, Hong Kong, China), Haoqian Wang (Tsinghua University, Shenzhen, China; Shenzhen Institute of Future Media Technology, Shenzhen, China), Qingyun Sun (Stanford University, CA, USA), Jun Xu (The Hong Kong Polytechnic University, Hong Kong, China), Qionghai Dai (Tsinghua University, Shenzhen, China; Shenzhen Institute of Future Media Technology, Shenzhen, China), and Lei Zhang (The Hong Kong Polytechnic University, Hong Kong, China)

Learning-Compression Algorithms for Neural Net Pruning 8532
Miguel A. Carreira-Perpinan (University of California, Merced) and Yerlan Idelbayev (University of California, Merced)

Large-Scale Distance Metric Learning with Uncertainty 8542
Qi Qian (Alibaba Group, WA, USA), Jiasheng Tang (Alibaba Group, WA, USA), Hao Li (Alibaba Group, WA, USA), Shenghuo Zhu (Alibaba Group, WA, USA), and Rong Jin (Alibaba Group, WA, USA)

Guide Me: Interacting with Deep Networks 8551
Christian Rupprecht (Technische Universitat München, Munich, Germany; Johns Hopkins University, Baltimore, MD, USA), Iro Laina (Technische Universitat München, Munich, Germany), Nassir Navab (Technische Universitat München, Munich, Germany), Gregory D. Hager (Johns Hopkins University, Baltimore, MD, USA; Technische Universitat München, Munich, Germany), and Federico Tombari (Technische Universitat München, Munich, Germany)

Art of Singular Vectors and Universal Adversarial Perturbations 8562
Ivan Oseledets (Skolkovo Institute of Science and Technology Institute of Numerical Mathematics RAS) and Valentin Khrulkov (Skolkovo Institute of Science and Technology)

Deflecting Adversarial Attacks with Pixel Deflection 8571
Aaditya Prakash (Brandeis University), Nick Moran (Brandeis University), Solomon Garber (Brandeis University), Antonella DiLillo (Brandeis University), and James Storer (Brandeis University)

MovieGraphs: Towards Understanding Human-Centric Situations from Videos 8581
Paul Vicol (University of Toronto; Vector Institute), Makarand Tapaswi (University of Toronto; Vector Institute), Lluís Castrejón (Montreal Institute for Learning Algorithms), and Sanja Fidler (University of Toronto; Vector Institute)

SemStyle: Learning to Generate Stylised Image Captions Using Unaligned Text 8591
Alexander Mathews (Australian National University; Data to Decision CRC), Lexing Xie (Australian National University; Data to Decision CRC), and Xuming He (ShanghaiTech University)

Benchmarking 6DOF Outdoor Visual Localization in Changing Conditions 8601
Torsten Sattler (ETH Zurich), Will Maddern (Oxford Robotics Institute, University of Oxford), Carl Toft (Chalmers University of Technology), Akihiko Torii (Tokyo Institute of Technology), Lars Hammarstrand (Chalmers University of Technology), Erik Stenborg (Chalmers University of Technology), Daniel Safari (Tokyo Institute of Technology; Technical University of Denmark), Masatoshi Okutomi (Tokyo Institute of Technology), Marc Pollefeys (ETH Zurich; Microsoft), Josef Sivic (Inria; CIIRC, CTU in Prague), Fredrik Kahl (Chalmers University of Technology; Centre for Mathematical Sciences, Lund University), and Tomas Pajdla (CIIRC, CTU in Prague)

iVQA: Inverse Visual Question Answering 8611
Feng Liu (Southeast University, China), Tao Xiang (Queen Mary University of London), Timothy M. Hospedales (3University of Edinburgh), Wankou Yang (Southeast University, China), and Changyin Sun (Southeast University, China)

Unsupervised Person Image Synthesis in Arbitrary Poses 8620
Albert Pumarola (Institut de Robòtica i Informàtica Industrial (CSIC-UPC), Barcelona, Spain), Antonio Agudo (Institut de Robòtica i Informàtica Industrial (CSIC-UPC), Barcelona, Spain), Alberto Sanfeliu (Institut de Robòtica i Informàtica Industrial (CSIC-UPC), Barcelona, Spain), and Francesc Moreno-Noguer (Institut de Robòtica i Informàtica Industrial (CSIC-UPC), Barcelona, Spain)

Oral and Spotlight 3-2C

O3-2C: Object Recognition & Scene Understanding V

Learning Descriptor Networks for 3D Shape Synthesis and Analysis 8629
Jianwen Xie (Hikvision Research Institute), Zilong Zheng (University of California, Los Angeles), Ruiqi Gao (University of California, Los Angeles), Wenguan Wang (University of California, Los Angeles; Beijing Institute of Technology), Song-Chun Zhu (University of California, Los Angeles), and Ying Nian Wu (University of California, Los Angeles)

Neural Kinematic Networks for Unsupervised Motion Retargetting 8639
Ruben Villegas (University of Michigan, Ann Arbor), Jimei Yang (Adobe Research), Duygu Ceylan (Adobe Research), and Honglak Lee (University of Michigan, Ann Arbor; Google Brain)

Group Consistent Similarity Learning via Deep CRF for Person Re-identification 8649
Dapeng Chen (CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong; School of Software Engineering, Xi'an Jiaotong University), Dan Xu (University of Trento), Hongsheng Li (CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong), Nicu Sebe (University of Trento), and Xiaogang Wang (CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong)

S3-2C

Learning Compositional Visual Concepts with Mutual Consistency 8659
Yunye Gong (Cornell University, Ithaca NY), Srikrishna Karanam (Siemens Corporate Technology, Princeton NJ), Ziyan Wu (Siemens Corporate Technology, Princeton NJ), Kuan-Chuan Peng (Siemens Corporate Technology, Princeton NJ), Jan Ernst (Siemens Corporate Technology, Princeton NJ), and Peter C. Doerschuk (Cornell University, Ithaca NY)

NestedNet: Learning Nested Sparse Structures in Deep Neural Networks 8669
Eunwoo Kim (Seoul National University, South Korea), Chanho Ahn (Seoul National University, South Korea), and Songhwai Oh (Seoul National University, South Korea)

Context Embedding Networks 8679
Kun Ho Kim (California Institute of Technology), Oisin Mac Aodha (California Institute of Technology), and Pietro Perona (California Institute of Technology)

Iterative Learning with Open-set Noisy Labels 8688
Yisen Wang (Tsinghua University, China; Georgia Institute of Technology), Weiyang Liu (Georgia Institute of Technology), Xingjun Ma (The University of Melbourne), James Bailey (The University of Melbourne), Hongyuan Zha (Georgia Institute of Technology), Le Song (Georgia Institute of Technology), and Shu-Tao Xia (Tsinghua University, China)

Learning Transferable Architectures for Scalable Image Recognition 8697
Barret Zoph (Google Brain), Vijay Vasudevan (Google Brain), Jonathon Shlens (Google Brain), and Quoc V. Le (Google Brain)

SBNet: Sparse Blocks Network for Fast Inference 8711
Mengye Ren (Uber Advanced Technologies Group; University of Toronto), Andrei Pokrovsky (Uber Advanced Technologies Group), Bin Yang (Uber Advanced Technologies Group; University of Toronto), and Raquel Urtasun (Uber Advanced Technologies Group; University of Toronto)

Language-Based Image Editing with Recurrent Attentive Models 8721
Jianbo Chen (University of California, Berkeley), Yelong Shen (Microsoft Research), Jianfeng Gao (Microsoft Research), Jingjing Liu (Microsoft Research), and Xiaodong Liu (Microsoft Research)

Net2Vec: Quantifying and Explaining How Concepts are Encoded by Filters in Deep Neural Networks 8730
Ruth Fong (University of Oxford) and Andrea Vedaldi (University of Oxford)

End-to-End Dense Video Captioning with Masked Transformer 8739
Luowei Zhou (University of Michigan), Yingbo Zhou (Salesforce Research), Jason J. Corso (University of Michigan), Richard Socher (Salesforce Research), and Caiming Xiong (Salesforce Research)

A Neural Multi-sequence Alignment TeCHnique (NeuMATCH) 8749
Pelin Dogan (ETH Zurich; Disney Research), Boyang Li (Liulishuo AI Lab), Leonid Sigal (University of British Columbia), and Markus Gross (ETH Zurich; Disney Research)

Path Aggregation Network for Instance Segmentation 8759
Shu Liu (The Chinese University of Hong Kong), Lu Qi (The Chinese University of Hong Kong), Haifang Qin (Peking University), Jianping Shi (SenseTime Research), and Jiaya Jia (The Chinese University of Hong Kong; YouTu Lab, Tencent)

The iNaturalist Species Classification and Detection Dataset 8769
Grant Van Horn (Caltech), Oisin Mac Aodha (Caltech), Yang Song (Google), Yin Cui (Cornell Tech), Chen Sun (Google), Alex Shepard (iNaturalist), Hartwig Adam (Google), Pietro Perona (Caltech), and Serge Belongie (Cornell Tech)

Multimodal Explanations: Justifying Decisions and Pointing to the Evidence 8779
Dong Huk Park (UC Berkeley), Lisa Anne Hendricks (UC Berkeley), Zeynep Akata (University of Amsterdam; MPI for Informatics), Anna Rohrbach (UC Berkeley; MPI for Informatics), Bernt Schiele (MPI for Informatics), Trevor Darrell (UC Berkeley), and Marcus Rohrbach (Facebook AI Research)

Oral and Spotlight 3-3A

O3-3A: Machine Learning for Computer Vision V

StarGAN: Unified Generative Adversarial Networks for Multi-domain Image-to-Image Translation 8789
Yunjey Choi (Korea University; Clova AI Research, NAVER Corp.), Minje Choi (Korea University; Clova AI Research, NAVER Corp.), Munyoung Kim (Clova AI Research, NAVER Corp.; The College of New Jersey), Jung-Woo Ha (Clova AI Research, NAVER Corp.), Sunghun Kim (Clova AI Research, NAVER Corp.; Hong Kong University of Science & Technology), and Jaegul Choo (Korea University; Clova AI Research, NAVER Corp.)

High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs 8798
Ting-Chun Wang (NVIDIA Corporation), Ming-Yu Liu (NVIDIA Corporation), Jun-Yan Zhu (UC Berkeley), Andrew Tao (NVIDIA Corporation), Jan Kautz (NVIDIA Corporation), and Bryan Catanzaro (NVIDIA Corporation)

Semi-Parametric Image Synthesis 8808
Xiaojuan Qi (CUHK), Qifeng Chen (Intel Labs), Jiaya Jia (CUHK), and Vladlen Koltun (Intel Labs)

S3-3A

BlockDrop: Dynamic Inference Paths in Residual Networks 8817
Zuxuan Wu (UMD), Tushar Nagarajan (UT Austin), Abhishek Kumar (IBM Research), Steven Rennie (Fusemachines Inc.), Larry S. Davis (UMD), Kristen Grauman (UT Austin), and Rogerio Feris (IBM Research)

Interpretable Convolutional Neural Networks 8827
Quanshi Zhang (University of California, Los Angeles), Ying Nian Wu (University of California, Los Angeles), and Song-Chun Zhu (University of California, Los Angeles)

Deep Cross-Media Knowledge Transfer 8837
Xin Huang (Peking University, Beijing, China) and Yuxin Peng (Peking University, Beijing, China)

Interleaved Structured Sparse Convolutional Neural Networks 8847
Guotian Xie (Sun Yat-Sen University; Guangdong Key Laboratory of Information Security Technology), Jingdong Wang (Microsoft Research), Ting Zhang (Microsoft Research), Jianhuang Lai (Sun Yat-Sen University; Guangdong Key Laboratory of Information Security Technology), Richang Hong (Hefei University of Technology), and Guo-Jun Qi (University of Central Florida)

A Variational U-Net for Conditional Appearance and Shape Generation 8857
Patrick Esser (Heidelberg University, Germany) and Ekaterina Sutter (Heidelberg University, Germany)

Detach and Adapt: Learning Cross-Domain Disentangled Deep Representation 8867
Yen-Cheng Liu (National Taiwan University, Taiwan), Yu-Ying Yeh (National Taiwan University, Taiwan), Tzu-Chien Fu (Northwestern University, USA), Sheng-De Wang (National Taiwan University, Taiwan), Wei-Chen Chiu (National Chiao Tung University, Taiwan), and Yu-Chiang Frank Wang (National Taiwan University, Taiwan)

Learning Deep Structured Active Contours End-to-End 8877
Lisa Zhang (University of Toronto, Canada), Min Bai (University of Toronto, Canada), Renjie Liao (University of Toronto, Canada), Raquel Urtasun (University of Toronto, Canada), Diego Marcos (University of Wageningen, Netherlands), Devis Tuia (University of Wageningen, Netherlands), and Benjamin Kellenberger (University of Wageningen, Netherlands)

Deep Learning Under Privileged Information Using Heteroscedastic Dropout 8886
John Lambert (Stanford University), Ozan Sener (Stanford University), and Silvio Savarese (Stanford University)

Smooth Neighbors on Teacher Graphs for Semi-Supervised Learning 8896
Yucen Luo (Tsinghua University, China), Jun Zhu (Tsinghua University, China), Mengxi Li (Tsinghua University, China), Yong Ren (Tsinghua University, China), and Bo Zhang (Tsinghua University, China)

Interpret Neural Networks by Identifying Critical Data Routing Paths 8906
Yulong Wang (Tsinghua University, China), Hang Su (Tsinghua University, China), Bo Zhang (Tsinghua University, China), and Xiaolin Hu (Tsinghua University, China)

Deep Spatio-Temporal Random Fields for Efficient Video Segmentation 8915
Siddhartha Chandra (INRIA GALEN, Ecole CentraleSupélec Paris), Camille Couprie (Facebook AI Research, Paris), and Iasonas Kokkinos (Facebook AI Research, Paris)

Customized Image Narrative Generation via Interactive Visual Question Generation and Answering 8925
Andrew Shin (The University of Tokyo), Yoshitaka Ushiku (The University of Tokyo), and Tatsuya Harada (The University of Tokyo; RIKEN)

Oral and Spotlight 3-3B

O3-3B: Image Motion & Tracking

PWC-Net: CNNs for Optical Flow Using Pyramid, Warping, and Cost Volume 8934
Deqing Sun (NVIDIA), Xiaodong Yang (NVIDIA), Ming-Yu Liu (NVIDIA), and Jan Kautz (NVIDIA)

Revisiting Deep Intrinsic Image Decompositions 8944
Qingnan Fan (Shandong University), Jiaolong Yang (Microsoft Research), Gang Hua (Microsoft Research), Baoquan Chen (Shenzhen Research Institute, Shandong University; Peking University), and David Wipf (Microsoft Research)

Multi-cell Detection and Classification Using a Generative Convolutional Model 8953
Florence Yellin (Johns Hopkins University), Benjamin D. Haeffele (Johns Hopkins University), Sophie Roth (IMEC), and Rene Vidal (Johns Hopkins University)

S3-3B

Learning Spatial-Aware Regressions for Visual Tracking 8962
Chong Sun (Dalian University of Technology, China; University of California, Merced, USA), Dong Wang (Dalian University of Technology, China), Huchuan Lu (Dalian University of Technology, China), and Ming-Hsuan Yang (University of California, Merced, USA)

High Performance Visual Tracking with Siamese Region Proposal Network 8971
Bo Li (SenseTime Group Limited; Beihang University), Junjie Yan (Tsinghua University, China), Wei Wu (SenseTime Group Limited), Zheng Zhu (SenseTime Group Limited; Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences), and Xiaolin Hu (Tsinghua University)

LiteFlowNet: A Lightweight Convolutional Neural Network for Optical Flow Estimation 8981
Tak-Wai Hui (CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong), Xiaoou Tang (CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong), and Chen Change Loy (CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong)

VITAL: VIsual Tracking via Adversarial Learning 8990
Yibing Song (Tencent AI Lab), Chao Ma (The University of Adelaide), Xiaohe Wu (Harbin Institute of Technology), Lijun Gong (Tencent), Linchao Bao (Tencent AI Lab), Wangmeng Zuo (Harbin Institute of Technology), Chunhua Shen (The University of Adelaide), Rynson W.H. Lau (City University of Hong Kong), and Ming-Hsuan Yang (University of California, Merced)

Super SloMo: High Quality Estimation of Multiple Intermediate Frames for Video Interpolation 9000
Huaizu Jiang (UMass Amherst), Deqing Sun (NVIDIA), Varan Jampani (NVIDIA), Ming-Hsuan Yang (NVIDIA; UC Merced), Erik Learned-Miller (UMass Amherst), and Jan Kautz (NVIDIA)

Real-World Repetition Estimation by Div, Grad and Curl 9009
Tom F. H. Runia (QUVA Deep Vision Lab, University of Amsterdam), Cees G. M. Snoek (QUVA Deep Vision Lab, University of Amsterdam), and Arnold W. M. Smeulders (QUVA Deep Vision Lab, University of Amsterdam)

Recurrent Pixel Embedding for Instance Grouping 9018
Shu Kong (University of California, Irvine) and Charless Fowlkes (University of California, Irvine)

Deep Unsupervised Saliency Detection: A Multiple Noisy Labeling Perspective 9029
Jing Zhang (Northwestern Polytechnical University, Xi'an, China; Australian National University, Canberra, Australia), Tong Zhang (Australian National University, Canberra, Australia; CSIRO, Canberra, Australia), Yuchao Daf (Northwestern Polytechnical University, Xi'an, China), Mehrtash Harandi (Australian National University, Canberra, Australia), Mehrtash Harandi (CSIRO, Canberra, Australia), and Richard Hartley (Australian National University, Canberra, Australia)

Learning Intrinsic Image Decomposition from Watching the World 9039
Zhengqi Li (Cornell University) and Noah Snavely (Cornell University)

TieNet: Text-Image Embedding Network for Common Thorax Disease Classification and Reporting in Chest X-Rays 9049
Xiaosong Wang (National Institutes of Health, Clinical Center), Yifan Peng (National Institutes of Health, Clinical Center), Le Lu (National Institutes of Health, Clinical Center), Zhiyong Lu (National Institutes of Health, Clinical Center), and Ronald M. Summers (National Institutes of Health, Clinical Center)

Generating Synthetic X-Ray Images of a Person from the Surface Geometry 9059
Brian Teixeira (Siemens Medical Solutions Inc., Princeton, NJ, USA), Vivek Singh (Siemens Medical Solutions Inc., Princeton, NJ, USA), Terrence Chen (Siemens Medical Solutions Inc., Princeton, NJ, USA), Kai Ma (Siemens Medical Solutions Inc., Princeton, NJ, USA), Birgi Tamersoy (Siemens Medical Solutions Inc., Princeton, NJ, USA), Yifan Wu (Temple University, Philadelphia, PA, USA), Elena Balashova (Princeton University, Princeton, NJ, USA), and Dorin Comaniciu (Siemens Medical Solutions Inc., Princeton, NJ, USA)

Gibson Env: Real-World Perception for Embodied Agents 9068
Fei Xia (Stanford University), Amir R. Zamir (Stanford University; University of California, Berkeley), Zhiyang He (Stanford University), Alexander Sax (Stanford University), Jitendra Malik (University of California, Berkeley), and Silvio Savarese (Stanford University)

Reinforcement Cutting-Agent Learning for Video Object Segmentation 9080
Junwei Han (Northwestern Polytechincal University), Le Yang (Northwestern Polytechincal University), Dingwen Zhang (Northwestern Polytechincal University; Xidian University), Xiaojun Chang (Carnegie Mellon University), and Xiaodan Liang (Carnegie Mellon University)

Oral and Spotlight 3-3C

O3-3C: Machine Learning for Computer Vision VI

Feature Space Transfer for Data Augmentation 9090
Bo Liu (University of California, San Diego), Xudong Wang (University of California, San Diego), Mandar Dixit (Microsoft), Roland Kwitt (University of Salzburg, Austria), and Nuno Vasconcelos (University of California, San Diego)

Analytic Expressions for Probabilistic Moments of PL-DNN with Gaussian Input 9099
Adel Bibi, Modar Alfadly, and Bernard Ghanem

Detail-Preserving Pooling in Deep Networks 9108
Faraz Saeedan (TU Darmstadt), Nicolas Weber (TU Darmstadt; NEC Laboratories Europe), Michael Goesele (TU Darmstadt; Oculus Research), and Stefan Roth (TU Darmstadt)

S3-3C

Rethinking Feature Distribution for Loss Functions in Image Classification 9117
Weitao Wan (Tsinghua University, Beijing, China), Yuanyi Zhong (Tsinghua University, Beijing, China; University of at Urbana-Champaign, Illinois, USA), Tianpeng Li (Tsinghua University, Beijing, China), and Jiansheng Chen (Tsinghua University, Beijing, China)

Shift: A Zero FLOP, Zero Parameter Alternative to Spatial Convolutions 9127
Bichen Wu (UC Berkeley), Alvin Wan (UC Berkeley), Xiangyu Yue (UC Berkeley), Peter Jin (UC Berkeley), Sicheng Zhao (UC Berkeley), Noah Golmant (UC Berkeley), Amir Gholaminejad (UC Berkeley), Joseph Gonzalez (UC Berkeley), and Kurt Keutzer (UC Berkeley)

Sketch-a-Classifier: Sketch-Based Photo Classifier Generation 9136
Conghui Hu (SketchX, Queen Mary University of London), Da Li (SketchX, Queen Mary University of London), Yi-Zhe Song (SketchX, Queen Mary University of London), Tao Xiang (SketchX, Queen Mary University of London), and Timothy M. Hospedales (SketchX, Queen Mary University of London; The University of Edinburgh)

Light Field Intrinsics with a Deep Encoder-Decoder Network 9145
Anna Alperovich (University of Konstanz Konstanz, Germany), Ole Johannsen (University of Konstanz Konstanz, Germany), Michael Strecke (University of Konstanz Konstanz, Germany), and Bastian Goldluecke (University of Konstanz Konstanz, Germany)

Learning Generative ConvNets via Multi-grid Modeling and Sampling 9155
Ruiqi Gao (University of California, Los Angeles, USA), Yang Lu (Amazon), Junpei Zhou (Zhejiang University, China), Song-Chun Zhu (University of California, Los Angeles, USA), and Ying Nian Wu (University of California, Los Angeles, USA)

Manifold Learning in Quotient Spaces 9165
Éloi Mehr (Sorbonne Université), André Lieutier (Sorbonne Université), Fernando Sanchez Bermudez (Sorbonne Université), Vincent Guitteny (Sorbonne Université), Nicolas Thome (Sorbonne Université), and Matthieu Cord (Sorbonne Université)

Learning Intelligent Dialogs for Bounding Box Annotation 9175
Jasper Uijlings (Google AI Perception), Ksenia Konyushkova (CVLab, EPFL), Christoph H. Lampert (IST Austria), and Vittorio Ferrari (Google AI Perception)

Boosting Adversarial Attacks with Momentum 9185
Yinpeng Dong (Tsinghua Lab of Brain and Intelligence; BNRist Lab; Tsinghua University, China), Fangzhou Liao (Tsinghua Lab of Brain and Intelligence; BNRist Lab; Tsinghua University, China), Tianyu Pang (Tsinghua Lab of Brain and Intelligence; BNRist Lab; Tsinghua University, China), Hang Su (Tsinghua Lab of Brain and Intelligence; BNRist Lab; Tsinghua University, China), Jun Zhu (Tsinghua Lab of Brain and Intelligence; BNRist Lab; Tsinghua University, China), Xiaolin Hu (Tsinghua Lab of Brain and Intelligence; BNRist Lab; Tsinghua University, China), and Jianguo Li (Intel Labs China)

NISP: Pruning Networks Using Neuron Importance Score Propagation 9194
Ruichi Yu (University of Maryland, College Park), Ang Li, Chun-Fu Chen (IBM T. J. Watson Research), Jui-Hsin Lai (JD.com), Vlad I. Morariu (Adobe Research), Xintong Han (University of Maryland, College Park), Mingfei Gao (University of Maryland, College Park), Ching-Yung Lin (Graphen.ai), and Larry S. Davis (University of Maryland, College Park)

PointGrid: A Deep Network for 3D Shape Understanding 9204
Truc Le (University of Missouri - Columbia) and Ye Duan (University of Missouri - Columbia)

Tell Me Where to Look: Guided Attention Inference Network 9215
Kunpeng Li (Northeastern University, Boston, MA), Ziyan Wu (Siemens Corporate Technology, Princeton, NJ), Kuan-Chuan Peng (Siemens Corporate Technology, Princeton, NJ), Jan Ernst (Siemens Corporate Technology, Princeton, NJ), and Yun Fu (Northeastern University, Boston, MA)

3D Semantic Segmentation with Submanifold Sparse Convolutional Networks 9224
Benjamin Graham (Facebook AI Research), Martin Engelcke (University of Oxford), and Laurens van der Maaten (Facebook AI Research)

TOM-Net: Learning Transparent Object Matting from a Single Image 9233
Guanying Chen (The University of Hong Kong, Hong Kong), Kai Han (The University of Hong Kong, Hong Kong), and Kwan-Yee K. Wong (The University of Hong Kong, Hong Kong)

Poster Session P3-2

Translating and Segmenting Multimodal Medical Volumes with Cycle- and Shape-Consistency Generative Adversarial Network 9242
Zizhao Zhang (University of Florida; Medical Imaging Technologies, Siemens Healthcare), Lin Yang (University of Florida), and Yefeng Zheng (Medical Imaging Technologies, Siemens Healthcare)

An Unsupervised Learning Model for Deformable Medical Image Registration 9252
Guha Balakrishnan (MIT), Amy Zhao (MIT), Mert R. Sabuncu (Cornell University), Adrian V. Dalca (MIT and MGH), and John Guttag (MIT)

Deep Lesion Graphs in the Wild: Relationship Learning and Organization of Significant Radiology Image Findings in a Diverse Large-Scale Lesion Database 9261
Ke Yan (National Institutes of Health, Clinical Center), Xiaosong Wang (National Institutes of Health, Clinical Center), Le Lu (National Institutes of Health, Clinical Center), Ling Zhang (National Institutes of Health, Clinical Center), Adam P. Harrison (National Institutes of Health, Clinical Center), Mohammadhadi Bagheri (National Institutes of Health, Clinical Center), and Ronald M. Summers (National Institutes of Health, Clinical Center)

Learning Distributions of Shape Trajectories from Longitudinal Datasets: A Hierarchical Model on a Manifold of Diffeomorphisms 9271
Alexandre Bone (Sorbonne Université), Olivier Colliot (Sorbonne Université), and Stanley Durrleman (Sorbonne Université)

CNN Driven Sparse Multi-level B-Spline Image Registration 9281
Pingge Jiang (Drexel University) and James A. Shackleford (Drexel University)

Anatomical Priors in Convolutional Networks for Unsupervised Biomedical Segmentation 9290
Adrian V. Dalca (MIT and MGH), John Guttag (MIT), and Mert R. Sabuncu (Cornell University)

3D Registration of Curves and Surfaces Using Local Differential Information 9300
Carolina Raposo (Institute of Systems and Robotics, University of Coimbra, Portugal; Perceive3D, Coimbra, Portugal) and Joao Barreto (Institute of Systems and Robotics, University of Coimbra, Portugal; Perceive3D, Coimbra, Portugal)

Weakly Supervised Learning of Single-Cell Feature Embeddings 9309
Juan C. Caicedo (Broad Institute of MIT and Harvard Cambridge, MA, USA), Claire McQuin (Broad Institute of MIT and Harvard Cambridge, MA, USA), Allen Goodman (Broad Institute of MIT and Harvard Cambridge, MA, USA), Shantanu Singh (Broad Institute of MIT and Harvard Cambridge, MA, USA), and Anne E. Carpenter (Broad Institute of MIT and Harvard Cambridge, MA, USA)

Guided Proofreading of Automatic Segmentations for Connectomics 9319
Daniel Haehn (Harvard University), Verena Kaynig (Harvard University), James Tompkin (Brown University), Jeff W. Lichtman (Harvard University), and Hanspeter Pfister (Harvard University)

Wide Compression: Tensor Ring Nets 9329
Vaneet Aggarwal (Purdue University), Wenlin Wang (Duke University), Brian Eriksson (Adobe), Yifan Sun (Technicolor Research), and Wenqi Wang (Purdue University)

Improvements to Context Based Self-Supervised Learning 9339
T. Nathan Mundhenk (Computational Engineering Division, Lawrence Livermore National Laboratory, Livermore, California), Daniel Ho, and Barry Y. Chen (Computational Engineering Division, Lawrence Livermore National Laboratory, Livermore, California)

Learning Structure and Strength of CNN Filters for Small Sample Size Training 9349
Rohit Keshari (IIIT-Delhi, India), Mayank Vatsa (IIIT-Delhi, India), Richa Singh (IIIT-Delhi, India), and Afzel Noore (Texas A&M University-Kingsville, USA)

Boosting Self-Supervised Learning via Knowledge Transfer 9359
Mehdi Noroozi (University of Bern), Ananth Vinjimoor (University of Maryland, Baltimore County), Paolo Favaro (University of Bern), and Hamed Pirsiavash (University of Maryland, Baltimore County)

The Power of Ensembles for Active Learning in Image Classification 9368
William H. Beluch (BCAI), Tim Genewein (BCAI), Andreas Nurnberger (University Magdeburg), and Jan M. Kohler (BCAI)

Learning Compact Recurrent Neural Networks with Block-Term Tensor Decomposition 9378
Jinmian Ye (University of Electronic Science and Technology of China), Linnan Wang (Brown University), Guangxi Li (University of Electronic Science and Technology of China), Di Chen (University of Electronic Science and Technology of China), Shandian Zhe (University of Utah), Xinqi Chu (Xjera Labs, Pte.Ltd), and Zenglin Xu (University of Electronic Science and Technology of China)

Spatially-Adaptive Filter Units for Deep Neural Networks 9388
Domen Tabemik (University of Ljubljana, Ljubljana, Slovenia), Matej Kristan (University of Ljubljana, Ljubljana, Slovenia), and Ales Leonardis (University of Ljubljana, Ljubljana, Slovenia; University of Birmingham, Birmingham)

SO-Net: Self-Organizing Network for Point Cloud Analysis 9397
Jiaxin Li (National University of Singapore), Ben M. Chen (National University of Singapore), and Gim Hee Lee (National University of Singapore)

SGAN: An Alternative Training of Generative Adversarial Networks 9407
Tatjana Chavdarova (Machine Learning group, Idiap Research Institute & École Polytechnique Fédérale de Lausanne) and François Fleuret (Machine Learning group, Idiap Research Institute & École Polytechnique Fédérale de Lausanne)

SketchyGAN: Towards Diverse and Realistic Sketch to Image Synthesis 9416
Wengling Chen (Georgia Institute of Technology) and James Hays (Georgia Institute of Technology, Argo AI)

Explicit Loss-Error-Aware Quantization for Low-Bit Deep Neural Networks 9426
Aojun Zhou (Intel Labs China), Anbang Yao (Intel Labs China), Kuan Wang (Tsinghua University, China), and Yurong Chen (Intel Labs China)

Towards Universal Representation for Unseen Action Recognition 9436
Yi Zhu (University of California, Merced), Yang Long (Open Lab, School of Computing, Newcastle University), Yu Guan (Open Lab, School of Computing, Newcastle University), Shawn Newsam (University of California, Merced), and Ling Shao (Inception Institute of Artificial Intelligence (IIAI), Abu Dhabi)

Deep Image Prior 9446
Victor Lempitsky (Skolkovo Institute of Science and Technology (Skoltech)), Andrea Vedaldi (University of Oxford), and Dmitry Ulyanov (Skolkovo Institute of Science and Technology, Yandex)

ST-GAN: Spatial Transformer Generative Adversarial Networks for Image Compositing 9455
Chen-Hsuan Lin (Carnegie Mellon University), Ersin Yumer (Adobe Research;), Oliver Wang (Adobe Research), Eli Shechtman (Adobe Research), and Simon Lucey (Carnegie Mellon University; Argo AI)

CartoonGAN: Generative Adversarial Networks for Photo Cartoonization 9465
Yang Chen (Tsinghua University, China), Yu-Kun Lai (Cardiff University), and Yong-Jin Liu (Tsinghua University, China)

Author Index